

01-17-01

JC06 Rec'd PCT/PTO 16 JAN 2001

PCT

FORM PTO-1390 (Modified) (REV 10-95)		U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE		ATTORNEY'S DOCKET NUMBER BB-1159	
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371				U.S. APPLICATION NO. (IF KNOWN, SEE 37 CFR) 09/744099	
INTERNATIONAL APPLICATION NO. PCT/US99/16354		INTERNATIONAL FILING DATE 20 JULY 1999 (20.07.99)		PRIORITY DATE CLAIMED 21 JULY 1998 (21.07.98)	
TITLE OF INVENTION CHORISMATE BIOSYNTHESIS ENZYMES					
APPLICANT(S) FOR DO/EO/US CAHOON, Rebecca E., et al.					
Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information					
<div>09/744099-01601</div> <div>1. <input checked="" type="checkbox"/> This is a FIRST submission of items concerning a filing under 35 U.S.C. 371.</div> <div>2. <input type="checkbox"/> This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.</div> <div>3. <input checked="" type="checkbox"/> This is an express request to being national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b)) and PCT Articles 22 and 39(1).</div> <div>4. <input checked="" type="checkbox"/> A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.</div> <div>5. <input checked="" type="checkbox"/> A copy of the International Application was filed (35 U.S.C. 371 (c) (2))<div>a. <input checked="" type="checkbox"/> is transmitted herewith (required only if not transmitted by the International Bureau.</div><div>b. <input type="checkbox"/> has been transmitted by the International Bureau.</div><div>c. <input type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US)</div></div> <div>6. <input type="checkbox"/> A translation of the International Application into English (35 U.S.C. 371 (c) (2)).</div> <div>7. <input checked="" type="checkbox"/> A copy of the International Search Report (PCT/ISA/210).</div> <div>8. <input checked="" type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371 (c) (3))<div>a. <input type="checkbox"/> are transmitted herewith (required only if not transmitted by the International Bureau).</div><div>b. <input type="checkbox"/> have been transmitted by the International Bureau.</div><div>c. <input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired.</div><div>d. <input checked="" type="checkbox"/> have not been made and will not be made.</div></div> <div>9. <input type="checkbox"/> A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371 (c)(3)).</div> <div>10. <input checked="" type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. 371 (c)(4)).</div> <div>11. <input checked="" type="checkbox"/> A copy of the International Preliminary Examination Report (PCT/IPEA/409)</div> <div>12. <input type="checkbox"/> A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371 (c)(5)).</div> <div>Items 13 to 18 below concern document(s) or information included :</div> <div>13. <input type="checkbox"/> An Information Disclosure Statement under 37 CFR 1.97 and 1.98.</div> <div>14. <input checked="" type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.</div> <div>15. <input type="checkbox"/> A FIRST preliminary amendment. A SECOND or SUBSEQUENT preliminary amendment.</div> <div>16. <input type="checkbox"/> A substitute specification.</div> <div>17. <input checked="" type="checkbox"/> A change of power of attorney and/or address letter.</div> <div>18. <input checked="" type="checkbox"/> Certificate of Mailing by Express Mail.</div> <div>19. <input type="checkbox"/> Other items or information:</div> <div><div>17. General Power of Attorney</div><div>18. Express Mailing Label No.: EL-031052246US</div><div>mailed 1/16/01</div></div>					

APPLICATION NO. (IF KNOWN, SEE 37 CFR) 09/744099		INTERNATIONAL APPLICATION NO. PCT/US99/16354		ATTORNEY'S DOCKET NUMBER BB-1159	
20. The following fees are submitted BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)) : <input checked="" type="checkbox"/> Search Report has been prepared by the EPO or JPO \$860.00 <input type="checkbox"/> International preliminary examination fee paid to USPTO (37 CFR 1.482) \$690.00 <input type="checkbox"/> No international preliminary examination fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO (37 CFR 1.445(a)(2)) \$760.00 <input type="checkbox"/> Neither international preliminary examination fee paid to USPTO (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO \$1000.00 <input type="checkbox"/> International preliminary examination fee paid to USPTO (37 CFR 1.482) And all claims satisfied provisions of PCT Article 33(2)-(4) \$ 100.00 ENTER APPROPRIATE BASIC FEE AMOUNT = \$860.00				CALCULATIONS PTO USE ONLY	
Surcharge of \$130.00 for furnishing the oath or declaration later than months from the earliest claimed priority date (37 CFR 1.492 (e)). <input type="checkbox"/> 20 <input type="checkbox"/> 30				\$0.00	
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total Claims	12 - 20 =	0 x	\$18.00	\$0.00	
Independent Claims	4 - 3 =	0 x	\$80.00	\$80.00	
Multiple Dependent Claims (check if applicable)			<input type="checkbox"/>	\$0.00	
TOTAL OF ABOVE CALCULATIONS =				\$80.00	
Reduction of 1/2 for filing by small entity, if applicable. Verified Small Entity Statement must also be filed (Note 37 CFR 1.9, 1.27, 1.28) (check if applicable).			<input type="checkbox"/>	\$0.00	
SUBTOTAL =				\$80.00	
Processing Fee of \$130.00 for furnishing the English translation later than months from the earliest claimed priority date (37 CFR 1.492 (f)). <input type="checkbox"/> 20 <input type="checkbox"/> 30				\$0.00	
TOTAL NATIONAL FEE =				\$940.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31) (check if applicable).			<input checked="" type="checkbox"/>	\$40.00	
TOTAL FEES ENCLOSED =				\$980.00	
				Amount to be : refunded	\$
				Charged	\$

- ☐ A check in the amount of _____ to cover the above fees enclosed.
- ☒ Please charge my Deposit Account No. **04-1928** in the amount of **\$980.00** to cover the above fees.
- ☒ The Commissioner is hereby authorized to charge any fees which may be required, or credit any overpayment to Deposit Account No. **04-1928** a duplicate copy of this sheet is enclosed.

NOTE : Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (CFR 1.37(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

LI, KENING
 E. I. DU PONT DE NEMOURS AND COMPANY
 Legal Patent Records Center
 1007 Market Street
 Wilmington, Delaware 19898
 United States of America

SIGNATURE

LI, KENING

NAME

44,872

REGISTRATION NUMBER

1/15/2001

DATE

TITLE

09/744099

CHORISMATE BIOSYNTHESIS ENZYMES

This application claims the benefit of U.S. Provisional Application No. 60/093,611, filed July 21, 1998.

FIELD OF THE INVENTION

This invention is in the field of plant molecular biology. More specifically, this invention pertains to nucleic acid fragments encoding enzymes involved in chorismate biosynthesis in plants and seeds.

BACKGROUND OF THE INVENTION

Chorismate biosynthesis involves the last few steps in the common pathway for the production of the aromatic amino acids phenylalanine, tyrosine and tryptophan. The first step in chorismate biosynthesis is performed by 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase. Two differentially expressed forms of this enzyme are present in plant tissues and are important regulators of the flux to aromatic amino acid biosynthesis. The next enzyme involved in this pathway is 3-dehydroquinate synthase which has been previously described in bacteria, fungi and yeast, (Millar, G. and Coggins, J.R. (1986) *FEBS Lett.* 200:11-17). This enzyme is encoded by the AroB gene in *E. coli* K12, its activity is dependent on NAD⁺ and divalent cations, and it is inhibited by the broad spectrum herbicide glyphosate. No plant 3-dehydroquinate synthase gene is known to date.

Manipulating either the amount or activity of this enzyme would afford a means to change the ratio of aromatic to non-aromatic amino acids in plants, including corn, rice, sorghum, soybean and wheat. This enzyme should also be useful for high throughput screening of compounds suitable for use as herbicides.

SUMMARY OF THE INVENTION

The instant invention relates to isolated nucleic acid fragments encoding 3-dehydroquinate synthase. Specifically, this invention concerns an isolated nucleic acid fragment encoding a 3-dehydroquinate synthase and an isolated nucleic acid fragment that is substantially similar to an isolated nucleic acid fragment encoding a 3-dehydroquinate synthase. In addition, this invention relates to a nucleic acid fragment that is complementary to the nucleic acid fragment encoding 3-dehydroquinate synthase.

An additional embodiment of the instant invention pertains to a polypeptide encoding all or a substantial portion of a 3-dehydroquinate synthase.

In another embodiment, the instant invention relates to a chimeric gene encoding a 3-dehydroquinate synthase, or to a chimeric gene that comprises a nucleic acid fragment that is complementary to a nucleic acid fragment encoding a 3-dehydroquinate synthase, operably linked to suitable regulatory sequences, wherein expression of the chimeric gene results in production of levels of the encoded protein in a transformed host cell that is altered (i.e., increased or decreased) from the level produced in an untransformed host cell.

In a further embodiment, the instant invention concerns a transformed host cell comprising in its genome a chimeric gene encoding a 3-dehydroquinate synthase, operably linked to suitable regulatory sequences. Expression of the chimeric gene results in production of altered levels of the encoded protein in the transformed host cell. The transformed host cell can be of eukaryotic or prokaryotic origin, and include cells derived from higher plants and microorganisms. The invention also includes transformed plants that arise from transformed host cells of higher plants, and seeds derived from such transformed plants.

An additional embodiment of the instant invention concerns a method of altering the level of expression of a 3-dehydroquinate synthase in a transformed host cell comprising: a) transforming a host cell with a chimeric gene comprising a nucleic acid fragment encoding a 3-dehydroquinate synthase; and b) growing the transformed host cell under conditions that are suitable for expression of the chimeric gene wherein expression of the chimeric gene results in production of altered levels of 3-dehydroquinate synthase in the transformed host cell.

An addition embodiment of the instant invention concerns a method for obtaining a nucleic acid fragment encoding all or a substantial portion of an amino acid sequence encoding a 3-dehydroquinate synthase.

A further embodiment of the instant invention is a method for evaluating at least one compound for its ability to inhibit the activity of a 3-dehydroquinate synthase, the method comprising the steps of: (a) transforming a host cell with a chimeric gene comprising a nucleic acid fragment encoding a 3-dehydroquinate synthase, operably linked to suitable regulatory sequences; (b) growing the transformed host cell under conditions that are suitable for expression of the chimeric gene wherein expression of the chimeric gene results in production of 3-dehydroquinate synthase in the transformed host cell; (c) optionally purifying the 3-dehydroquinate synthase expressed by the transformed host cell; (d) treating the 3-dehydroquinate synthase with a compound to be tested; and (e) comparing the activity of the 3-dehydroquinate synthase that has been treated with a test compound to the activity of an untreated 3-dehydroquinate synthase, thereby selecting compounds with potential for inhibitory activity.

BRIEF DESCRIPTION OF THE DRAWING AND SEQUENCE DESCRIPTIONS

The invention can be more fully understood from the following detailed description and the accompanying drawing and Sequence Listing which form a part of this application.

Figure 1 depicts the amino acid sequence alignment between the 3-dehydroquinate synthase from *Escherichia coli* (NCBI General Identifier No. 114181; SEQ ID NO:9), the amino acid sequence encoded by the instant corn contig assembled from cDNA clones cr1n.pk0113.g3cbn2.pk0047.e10 and cca.pk0019.e12 (SEQ ID NO:2), the amino acid

sequence encoded by the instant rice cDNA clone rls72.pk0035.a10 (SEQ ID NO:4), the amino acid sequence encoded by the instant soybean cDNA clone se3.pk0029.f9 (SEQ ID NO:6) and the amino acid sequence encoded by the instant wheat cDNA clone wr1.pk0011.d5 (SEQ ID NO:8). The top row indicates with asterisks (*) the amino acids conserved among all sequences and with plus signs (+) the amino acids conserved only among the plant sequences. Dashes are used by the program to maximize the alignment of the sequences.

Table 1 lists the polypeptides that are described herein, the designation of the cDNA clones that comprise the nucleic acid fragments encoding polypeptides representing all or a substantial portion of these polypeptides, and the corresponding identifier (SEQ ID NO:) as used in the attached Sequence Listing. The sequence descriptions and Sequence Listing attached hereto comply with the rules governing nucleotide and/or amino acid sequence disclosures in patent applications as set forth in 37 C.F.R. §1.821-1.825.

TABLE 1

3-Dehydroquinate Synthase

Plant	Clone Designation	SEQ ID NO:	
		(Nucleotide)	(Amino Acid)
Corn	Contig of:	1	2
	cbn2.pk0047.e10		
	cca.pk0019.e12		
Rice	rls72.pk0035.a10	3	4
Soybean	se3.pk0029.f9	5	6
Wheat	wr1.pk0011.d5	7	8

The Sequence Listing contains the one letter code for nucleotide sequence characters and the three letter codes for amino acids as defined in conformity with the IUPAC-IUBMB standards described in *Nucleic Acids Research* 13:3021-3030 (1985) and in the *Biochemical Journal* 219 (No. 2):345-373 (1984) which are herein incorporated by reference. The symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. §1.822.

DETAILED DESCRIPTION OF THE INVENTION

In the context of this disclosure, a number of terms shall be utilized. As used herein, a "nucleic acid fragment" is a polymer of RNA or DNA that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. A nucleic acid fragment in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA.

As used herein, "contig" refers to a nucleotide sequence that is assembled from two or more constituent nucleotide sequences that share common or overlapping regions of

sequence homology. For example, the nucleotide sequences of two or more nucleic acid fragments can be compared and aligned in order to identify common or overlapping sequences. Where common or overlapping sequences exist between two or more nucleic acid fragments, the sequences (and thus their corresponding nucleic acid fragments) can be assembled into a single contiguous nucleotide sequence.

As used herein, "substantially similar" refers to nucleic acid fragments wherein changes in one or more nucleotide bases results in substitution of one or more amino acids, but do not affect the functional properties of the polypeptide encoded by the nucleotide sequence. "Substantially similar" also refers to nucleic acid fragments wherein changes in one or more nucleotide bases does not affect the ability of the nucleic acid fragment to mediate alteration of gene expression by gene silencing through for example antisense or co-suppression technology. "Substantially similar" also refers to modifications of the nucleic acid fragments of the instant invention such as deletion or insertion of one or more nucleotides that do not substantially affect the functional properties of the resulting transcript vis-à-vis the ability to mediate gene silencing or alteration of the functional properties of the resulting protein molecule. It is therefore understood that the invention encompasses more than the specific exemplary nucleotide or amino acid sequences and includes functional equivalents thereof.

For example, it is well known in the art that antisense suppression and co-suppression of gene expression may be accomplished using nucleic acid fragments representing less than the entire coding region of a gene, and by nucleic acid fragments that do not share 100% sequence identity with the gene to be suppressed. Moreover, alterations in a nucleic acid fragment which result in the production of a chemically equivalent amino acid at a given site, but do not effect the functional properties of the encoded polypeptide, are well known in the art. Thus, a codon for the amino acid alanine, a hydrophobic amino acid, may be substituted by a codon encoding another less hydrophobic residue, such as glycine, or a more hydrophobic residue, such as valine, leucine, or isoleucine. Similarly, changes which result in substitution of one negatively charged residue for another, such as aspartic acid for glutamic acid, or one positively charged residue for another, such as lysine for arginine, can also be expected to produce a functionally equivalent product. Nucleotide changes which result in alteration of the N-terminal and C-terminal portions of the polypeptide molecule would also not be expected to alter the activity of the polypeptide. Each of the proposed modifications is well within the routine skill in the art, as is determination of retention of biological activity of the encoded products.

Moreover, substantially similar nucleic acid fragments may also be characterized by their ability to hybridize. Estimates of such homology are provided by either DNA-DNA or DNA-RNA hybridization under conditions of stringency as is well understood by those skilled in the art (Hames and Higgins, Eds. (1985) *Nucleic Acid Hybridisation*, IRL Press,

Oxford, U.K.). Stringency conditions can be adjusted to screen for moderately similar fragments, such as homologous sequences from distantly related organisms, to highly similar fragments, such as genes that duplicate functional enzymes from closely related organisms.

Post-hybridization washes determine stringency conditions. One set of preferred conditions

- 5 uses a series of washes starting with 6X SSC, 0.5% SDS at room temperature for 15 min, then repeated with 2X SSC, 0.5% SDS at 45°C for 30 min, and then repeated twice with 0.2X SSC, 0.5% SDS at 50°C for 30 min. A more preferred set of stringent conditions uses higher temperatures in which the washes are identical to those above except for the temperature of the final two 30 min washes in 0.2X SSC, 0.5% SDS was increased to 60°C.
- 10 Another preferred set of highly stringent conditions uses two final washes in 0.1X SSC, 0.1% SDS at 65°C.

Substantially similar nucleic acid fragments of the instant invention may also be characterized by the percent identity of the amino acid sequences that they encode to the amino acid sequences disclosed herein, as determined by algorithms commonly employed by

- 15 those skilled in this art. Preferred are those nucleic acid fragments whose nucleotide sequences encode amino acid sequences that are 80% identical to the amino acid sequences reported herein. More preferred nucleic acid fragments encode amino acid sequences that are 90% identical to the amino acid sequences reported herein. Most preferred are nucleic acid fragments that encode amino acid sequences that are 95% identical to the amino acid sequences reported herein. Sequence alignments and percent identity calculations were performed using the Megalign program of the LASARGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences was performed using the Clustal method of alignment (Higgins and Sharp (1989) *CABIOS* 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default
- 20 parameters for pairwise alignments using the Clustal method were KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5.

A "substantial portion" of an amino acid or nucleotide sequence comprises an amino acid or a nucleotide sequence that is sufficient to afford putative identification of the protein or gene that the amino acid or nucleotide sequence comprises. Amino acid and nucleotide

- 30 sequences can be evaluated either manually by one skilled in the art, or by using computer-based sequence comparison and identification tools that employ algorithms such as BLAST (Basic Local Alignment Search Tool; Altschul et al. (1993) *J. Mol. Biol.* 215:403-410; see also www.ncbi.nlm.nih.gov/BLAST/). In general, a sequence of ten or more contiguous amino acids or thirty or more contiguous nucleotides is necessary in order to putatively
- 35 identify a polypeptide or nucleic acid sequence as homologous to a known protein or gene. Moreover, with respect to nucleotide sequences, gene-specific oligonucleotide probes comprising 30 or more contiguous nucleotides may be used in sequence-dependent methods of gene identification (e.g., Southern hybridization) and isolation (e.g., *in situ* hybridization

of bacterial colonies or bacteriophage plaques). In addition, short oligonucleotides of 12 or more nucleotides may be used as amplification primers in PCR in order to obtain a particular nucleic acid fragment comprising the primers. Accordingly, a "substantial portion" of a nucleotide sequence comprises a nucleotide sequence that will afford specific identification and/or isolation of a nucleic acid fragment comprising the sequence. The instant specification teaches amino acid and nucleotide sequences encoding polypeptides that comprise one or more particular plant proteins. The skilled artisan, having the benefit of the sequences as reported herein, may now use all or a substantial portion of the disclosed sequences for purposes known to those skilled in this art. Accordingly, the instant invention comprises the complete sequences as reported in the accompanying Sequence Listing, as well as substantial portions of those sequences as defined above.

"Codon degeneracy" refers to divergence in the genetic code permitting variation of the nucleotide sequence without effecting the amino acid sequence of an encoded polypeptide. Accordingly, the instant invention relates to any nucleic acid fragment comprising a nucleotide sequence that encodes all or a substantial portion of the amino acid sequences set forth herein. The skilled artisan is well aware of the "codon-bias" exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. Therefore, when synthesizing a nucleic acid fragment for improved expression in a host cell, it is desirable to design the nucleic acid fragment such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

"Synthetic nucleic acid fragments" can be assembled from oligonucleotide building blocks that are chemically synthesized using procedures known to those skilled in the art. These building blocks are ligated and annealed to form larger nucleic acid fragments which may then be enzymatically assembled to construct the entire desired nucleic acid fragment. "Chemically synthesized", as related to nucleic acid fragment, means that the component nucleotides were assembled *in vitro*. Manual chemical synthesis of nucleic acid fragments may be accomplished using well established procedures, or automated chemical synthesis can be performed using one of a number of commercially available machines. Accordingly, the nucleic acid fragments can be tailored for optimal gene expression based on optimization of nucleotide sequence to reflect the codon bias of the host cell. The skilled artisan appreciates the likelihood of successful gene expression if codon usage is biased towards those codons favored by the host. Determination of preferred codons can be based on a survey of genes derived from the host cell where sequence information is available.

"Gene" refers to a nucleic acid fragment that expresses a specific protein, including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. "Native gene" refers to a gene as found in nature with its own regulatory sequences. "Chimeric gene" refers any gene that is not a native gene, comprising regulatory and coding sequences that are not found together in nature.

Accordingly, a chimeric gene may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature.

"Endogenous gene" refers to a native gene in its natural location in the genome of an

5 organism. A "foreign" gene refers to a gene not normally found in the host organism, but that is introduced into the host organism by gene transfer. Foreign genes can comprise native genes inserted into a non-native organism, or chimeric genes. A "transgene" is a gene that has been introduced into the genome by a transformation procedure.

10 "Coding sequence" refers to a nucleotide sequence that codes for a specific amino acid sequence. "Regulatory sequences" refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include promoters, translation leader sequences, introns, and polyadenylation recognition sequences.

15 "Promoter" refers to a nucleotide sequence capable of controlling the expression of a coding sequence or functional RNA. In general, a coding sequence is located 3' to a promoter sequence. The promoter sequence consists of proximal and more distal upstream elements, the latter elements often referred to as enhancers. Accordingly, an "enhancer" is a nucleotide sequence which can stimulate promoter activity and may be an innate element of
20 the promoter or a heterologous element inserted to enhance the level or tissue-specificity of a promoter. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic nucleotide segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at
25 different stages of development, or in response to different environmental conditions. Promoters which cause a nucleic acid fragment to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". New promoters of various types useful in plant cells are constantly being discovered; numerous examples may be found in the compilation by Okamuro and Goldberg (1989) *Biochemistry of Plants* 15:1-82.
30 It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, nucleic acid fragments of different lengths may have identical promoter activity.

The "translation leader sequence" refers to a nucleotide sequence located between the promoter sequence of a gene and the coding sequence. The translation leader sequence is
35 present in the fully processed mRNA upstream of the translation start sequence. The translation leader sequence may affect processing of the primary transcript to mRNA, mRNA stability or translation efficiency. Examples of translation leader sequences have been described (Turner and Foster (1995) *Molecular Biotechnology* 3:225).

The "3' non-coding sequences" refer to nucleotide sequences located downstream of a coding sequence and include polyadenylation recognition sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. The use of different 3' non-coding sequences is exemplified by Ingelbrecht et al. (1989) *Plant Cell* 1:671-680.

"RNA transcript" refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript or it may be a RNA sequence derived from posttranscriptional processing of the primary transcript and is referred to as the mature RNA. "Messenger RNA (mRNA)" refers to the RNA that is without introns and that can be translated into polypeptide by the cell. "cDNA" refers to a double-stranded DNA that is complementary to and derived from mRNA. "Sense" RNA refers to an RNA transcript that includes the mRNA and so can be translated into a polypeptide by the cell. "Antisense RNA" refers to an RNA transcript that is complementary to all or part of a target primary transcript or mRNA and that blocks the expression of a target gene (see U.S. Patent No. 5,107,065, incorporated herein by reference). The complementarity of an antisense RNA may be with any part of the specific nucleotide sequence, i.e., at the 5' non-coding sequence, 3' non-coding sequence, introns, or the coding sequence. "Functional RNA" refers to sense RNA, antisense RNA, ribozyme RNA, or other RNA that may not be translated but yet has an effect on cellular processes.

The term "operably linked" refers to the association of two or more nucleic acid fragments on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence when it is capable of affecting the expression of that coding sequence (i.e., that the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation.

The term "expression", as used herein, refers to the transcription and stable accumulation of sense (mRNA) or antisense RNA derived from the nucleic acid fragment of the invention. Expression may also refer to translation of mRNA into a polypeptide. "Antisense inhibition" refers to the production of antisense RNA transcripts capable of suppressing the expression of the target protein. "Overexpression" refers to the production of a gene product in transgenic organisms that exceeds levels of production in normal or non-transformed organisms. "Co-suppression" refers to the production of sense RNA transcripts capable of suppressing the expression of identical or substantially similar foreign or endogenous genes (U.S. Patent No. 5,231,020, incorporated herein by reference).

"Altered levels" refers to the production of gene product(s) in transgenic organisms in amounts or proportions that differ from that of normal or non-transformed organisms.

“Mature” protein refers to a post-translationally processed polypeptide; i.e., one from which any pre- or propeptides present in the primary translation product have been removed. “Precursor” protein refers to the primary product of translation of mRNA; i.e., with pre- and propeptides still present. Pre- and propeptides may be but are not limited to intracellular localization signals.

“Transformation” refers to the transfer of a nucleic acid fragment into the genome of a host organism, resulting in genetically stable inheritance. Host organisms containing the transformed nucleic acid fragments are referred to as “transgenic” organisms. Examples of methods of plant transformation include *Agrobacterium*-mediated transformation (De Blaere et al. (1987) *Meth. Enzymol.* 143:277) and particle-accelerated or “gene gun” transformation technology (Klein et al. (1987) *Nature (London)* 327:70-73; U.S. Patent No. 4,945,050, incorporated herein by reference).

Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described more fully in Sambrook et al. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1989 (hereinafter “Maniatis”).

Nucleic acid fragments encoding at least a portion of several 3-dehydroquinate synthases have been isolated and identified by comparison of random plant cDNA sequences to public databases containing nucleotide and protein sequences using the BLAST algorithms well known to those skilled in the art. The nucleic acid fragments of the instant invention may be used to isolate cDNAs and genes encoding homologous proteins from the same or other plant species. Isolation of homologous genes using sequence-dependent protocols is well known in the art. Examples of sequence-dependent protocols include, but are not limited to, methods of nucleic acid hybridization, and methods of DNA and RNA amplification as exemplified by various uses of nucleic acid amplification technologies (e.g., polymerase chain reaction, ligase chain reaction).

For example, genes encoding other 3-dehydroquinate synthases, either as cDNAs or genomic DNAs, could be isolated directly by using all or a portion of the instant nucleic acid fragments as DNA hybridization probes to screen libraries from any desired plant employing methodology well known to those skilled in the art. Specific oligonucleotide probes based upon the instant nucleic acid sequences can be designed and synthesized by methods known in the art (Maniatis). Moreover, the entire sequences can be used directly to synthesize DNA probes by methods known to the skilled artisan such as random primer DNA labeling, nick translation, or end-labeling techniques, or RNA probes using available *in vitro* transcription systems. In addition, specific primers can be designed and used to amplify a part or all of the instant sequences. The resulting amplification products can be labeled directly during amplification reactions or labeled after amplification reactions, and used as

probes to isolate full length cDNA or genomic fragments under conditions of appropriate stringency.

In addition, two short segments of the instant nucleic acid fragments may be used in polymerase chain reaction protocols to amplify longer nucleic acid fragments encoding homologous genes from DNA or RNA. The polymerase chain reaction may also be performed on a library of cloned nucleic acid fragments wherein the sequence of one primer is derived from the instant nucleic acid fragments, and the sequence of the other primer takes advantage of the presence of the polyadenylic acid tracts to the 3' end of the mRNA precursor encoding plant genes. Alternatively, the second primer sequence may be based upon sequences derived from the cloning vector. For example, the skilled artisan can follow the RACE protocol (Frohman et al. (1988) *Proc. Natl. Acad. Sci. USA* 85:8998) to generate cDNAs by using PCR to amplify copies of the region between a single point in the transcript and the 3' or 5' end. Primers oriented in the 3' and 5' directions can be designed from the instant sequences. Using commercially available 3' RACE or 5' RACE systems (BRL), specific 3' or 5' cDNA fragments can be isolated (Ohara et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:5673; Loh et al. (1989) *Science* 243:217). Products generated by the 3' and 5' RACE procedures can be combined to generate full-length cDNAs (Frohman and Martin (1989) *Techniques* 1:165).

Availability of the instant nucleotide and deduced amino acid sequences facilitates immunological screening of cDNA expression libraries. Synthetic peptides representing portions of the instant amino acid sequences may be synthesized. These peptides can be used to immunize animals to produce polyclonal or monoclonal antibodies with specificity for peptides or proteins comprising the amino acid sequences. These antibodies can be then be used to screen cDNA expression libraries to isolate full-length cDNA clones of interest (Lerner (1984) *Adv. Immunol.* 36:1; Maniatis).

The nucleic acid fragments of the instant invention may be used to create transgenic plants in which the disclosed polypeptides are present at higher or lower levels than normal or in cell types or developmental stages in which they are not normally found. This would have the effect of altering the ratio of aromatic to non-aromatic amino acids in those cells. This enzyme may also be used for the high throughput screening of compounds suitable for use as herbicides.

Overexpression of the proteins of the instant invention may be accomplished by first constructing a chimeric gene in which the coding region is operably linked to a promoter capable of directing expression of a gene in the desired tissues at the desired stage of development. For reasons of convenience, the chimeric gene may comprise promoter sequences and translation leader sequences derived from the same genes. 3' Non-coding sequences encoding transcription termination signals may also be provided. The instant chimeric gene may also comprise one or more introns in order to facilitate gene expression.

Plasmid vectors comprising the instant chimeric gene can then constructed. The choice of plasmid vector is dependent upon the method that will be used to transform host plants. The skilled artisan is well aware of the genetic elements that must be present on the plasmid vector in order to successfully transform, select and propagate host cells containing the chimeric gene. The skilled artisan will also recognize that different independent transformation events will result in different levels and patterns of expression (Jones et al. (1985) *EMBO J.* 4:2411-2418; De Almeida et al. (1989) *Mol. Gen. Genetics* 218:78-86), and thus that multiple events must be screened in order to obtain lines displaying the desired expression level and pattern. Such screening may be accomplished by Southern analysis of DNA, Northern analysis of mRNA expression, Western analysis of protein expression, or phenotypic analysis.

It may also be desirable to reduce or eliminate expression of genes encoding the instant polypeptide in plants for some applications. In order to accomplish this, a chimeric gene designed for co-suppression of the instant polypeptide can be constructed by linking a gene or gene fragment encoding that polypeptide to plant promoter sequences. Alternatively, a chimeric gene designed to express antisense RNA for all or part of the instant nucleic acid fragment can be constructed by linking the gene or gene fragment in reverse orientation to plant promoter sequences. Either the co-suppression or antisense chimeric genes could be introduced into plants via transformation wherein expression of the corresponding endogenous genes are reduced or eliminated.

Molecular genetic solutions to the generation of plants with altered gene expression have a decided advantage over more traditional plant breeding approaches. Changes in plant phenotypes can be produced by specifically inhibiting expression of one or more genes by antisense inhibition or cosuppression (U. S. Patent Nos. 5,190,931, 5,107,065 and 5,283,323). An antisense or cosuppression construct would act as a dominant negative regulator of gene activity. While conventional mutations can yield negative regulation of gene activity these effects are most likely recessive. The dominant negative regulation available with a transgenic approach may be advantageous from a breeding perspective. In addition, the ability to restrict the expression of specific phenotype to the reproductive tissues of the plant by the use of tissue specific promoters may confer agronomic advantages relative to conventional mutations which may have an effect in all tissues in which a mutant gene is ordinarily expressed.

The person skilled in the art will know that special considerations are associated with the use of antisense or cosuppression technologies in order to reduce expression of particular genes. For example, the proper level of expression of sense or antisense genes may require the use of different chimeric genes utilizing different regulatory elements known to the skilled artisan. Once transgenic plants are obtained by one of the methods described above, it will be necessary to screen individual transgenics for those that most effectively display

the desired phenotype. Accordingly, the skilled artisan will develop methods for screening large numbers of transformants. The nature of these screens will generally be chosen on practical grounds, and is not an inherent part of the invention. For example, one can screen by looking for changes in gene expression by using antibodies specific for the protein encoded by the gene being suppressed, or one could establish assays that specifically measure enzyme activity. A preferred method will be one which allows large numbers of samples to be processed rapidly, since it will be expected that a large number of transformants will be negative for the desired phenotype.

The instant polypeptide (or portions thereof) may be produced in heterologous host cells, particularly in the cells of microbial hosts, and can be used to prepare antibodies to these proteins by methods well known to those skilled in the art. The antibodies are useful for detecting the polypeptide of the instant invention *in situ* in cells or *in vitro* in cell extracts. Preferred heterologous host cells for production of the instant polypeptide are microbial hosts. Microbial expression systems and expression vectors containing regulatory sequences that direct high level expression of foreign proteins are well known to those skilled in the art. Any of these could be used to construct a chimeric gene for production of the instant polypeptide. This chimeric gene could then be introduced into appropriate microorganisms via transformation to provide high level expression of the encoded 3-dehydroquinate synthase. An example of a vector for high level expression of the instant polypeptide in a bacterial host is provided (Example 6).

Additionally, the instant polypeptide can be used as a target to facilitate design and/or identification of inhibitors of those enzymes that may be useful as herbicides. This is desirable because the polypeptide described herein catalyzes a step in aromatic amino acid biosynthesis. Accordingly, inhibition of the activity of one or more of the enzymes described herein could lead to inhibition plant growth. Thus, the instant polypeptide could be appropriate for new herbicide discovery and design.

All or a substantial portion of the nucleic acid fragments of the instant invention may also be used as probes for genetically and physically mapping the genes that they are a part of, and as markers for traits linked to those genes. Such information may be useful in plant breeding in order to develop lines with desired phenotypes. For example, the instant nucleic acid fragments may be used as restriction fragment length polymorphism (RFLP) markers. Southern blots (Maniatis) of restriction-digested plant genomic DNA may be probed with the nucleic acid fragments of the instant invention. The resulting banding patterns may then be subjected to genetic analyses using computer programs such as MapMaker (Lander et al. (1987) *Genomics* 1:174-181) in order to construct a genetic map. In addition, the nucleic acid fragments of the instant invention may be used to probe Southern blots containing restriction endonuclease-treated genomic DNAs of a set of individuals representing parent and progeny of a defined genetic cross. Segregation of the DNA polymorphisms is noted

and used to calculate the position of the instant nucleic acid sequence in the genetic map previously obtained using this population (Botstein et al. (1980) *Am. J. Hum. Genet.* 32:314-331).

The production and use of plant gene-derived probes for use in genetic mapping is described in Bernatzky and Tanksley (1986) *Plant Mol. Biol. Reporter* 4(1):37-41. Numerous publications describe genetic mapping of specific cDNA clones using the methodology outlined above or variations thereof. For example, F2 intercross populations, backcross populations, randomly mated populations, near isogenic lines, and other sets of individuals may be used for mapping. Such methodologies are well known to those skilled in the art.

Nucleic acid probes derived from the instant nucleic acid sequences may also be used for physical mapping (i.e., placement of sequences on physical maps; see Hoheisel et al. In: *Nonmammalian Genomic Analysis: A Practical Guide*, Academic press 1996, pp. 319-346, and references cited therein).

In another embodiment, nucleic acid probes derived from the instant nucleic acid sequences may be used in direct fluorescence *in situ* hybridization (FISH) mapping (Trask (1991) *Trends Genet.* 7:149-154). Although current methods of FISH mapping favor use of large clones (several to several hundred KB; see Laan et al. (1995) *Genome Research* 5:13-20), improvements in sensitivity may allow performance of FISH mapping using shorter probes.

A variety of nucleic acid amplification-based methods of genetic and physical mapping may be carried out using the instant nucleic acid sequences. Examples include allele-specific amplification (Kazazian (1989) *J. Lab. Clin. Med.* 114(2):95-96), polymorphism of PCR-amplified fragments (CAPS; Sheffield et al. (1993) *Genomics* 16:325-332), allele-specific ligation (Landegren et al. (1988) *Science* 241:1077-1080), nucleotide extension reactions (Sokolov (1990) *Nucleic Acid Res.* 18:3671), Radiation Hybrid Mapping (Walter et al. (1997) *Nature Genetics* 7:22-28) and Happy Mapping (Dear and Cook (1989) *Nucleic Acid Res.* 17:6795-6807). For these methods, the sequence of a nucleic acid fragment is used to design and produce primer pairs for use in the amplification reaction or in primer extension reactions. The design of such primers is well known to those skilled in the art. In methods employing PCR-based genetic mapping, it may be necessary to identify DNA sequence differences between the parents of the mapping cross in the region corresponding to the instant nucleic acid sequence. This, however, is generally not necessary for mapping methods.

Loss of function mutant phenotypes may be identified for the instant cDNA clones either by targeted gene disruption protocols or by identifying specific mutants for these genes contained in a maize population carrying mutations in all possible genes (Ballinger and Benzer (1989) *Proc. Natl. Acad. Sci USA* 86:9402; Koes et al. (1995) *Proc. Natl. Acad.*

Sci USA 92:8149; Bensen et al. (1995) *Plant Cell* 7:75). The latter approach may be accomplished in two ways. First, short segments of the instant nucleic acid fragments may be used in polymerase chain reaction protocols in conjunction with a mutation tag sequence primer on DNAs prepared from a population of plants in which Mutator transposons or some other mutation-causing DNA element has been introduced (see Bensen, *supra*). The amplification of a specific DNA fragment with these primers indicates the insertion of the mutation tag element in or near the plant gene encoding the instant polypeptide. Alternatively, the instant nucleic acid fragment may be used as a hybridization probe against PCR amplification products generated from the mutation population using the mutation tag sequence primer in conjunction with an arbitrary genomic site primer, such as that for a restriction enzyme site-anchored synthetic adaptor. With either method, a plant containing a mutation in the endogenous gene encoding the instant polypeptide can be identified and obtained. This mutant plant can then be used to determine or confirm the natural function of the instant polypeptide disclosed herein.

EXAMPLES

The present invention is further defined in the following Examples, in which all parts and percentages are by weight and degrees are Celsius, unless otherwise stated. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions.

EXAMPLE 1

Composition of cDNA Libraries; Isolation and Sequencing of cDNA Clones

cDNA libraries representing mRNAs from various corn, rice, soybean and wheat tissues were prepared. The characteristics of the libraries are described below.

TABLE 2

cDNA Libraries from Corn, Rice, Soybean and Wheat

Library	Tissue	Clone
cbn2	Corn Developing Kernel Two Days After Pollination	cbn2.pk0047.e10
cca	Corn Callus Type II Tissue, Undifferentiated, Highly Transformable	cca.pk0019.e12
rls72	Rice Leaf 15 Days After Germination, 72 Hours After Infection of Strain <i>Magaporthe grisea</i> 4360-R-67 (AVR2-YAMO); Susceptible	rls72.pk0035.a10
se3	Soybean Embryo, 17 Days After Flowering	se3.pk0029.f9
wr1	Wheat Root From 7 Day Old Seedling	wr1.pk0011.d5

cDNA libraries may be prepared by any one of many methods available. For example, the cDNAs may be introduced into plasmid vectors by first preparing the cDNA libraries in Uni-ZAP* XR vectors according to the manufacturer's protocol (Stratagene Cloning Systems, La Jolla, CA). The Uni-ZAP* XR libraries are converted into plasmid libraries according to the protocol provided by Stratagene. Upon conversion, cDNA inserts will be contained in the plasmid vector pBluescript. In addition, the cDNAs may be introduced directly into precut Bluescript II SK(+) vectors (Stratagene) using T4 DNA ligase (New England Biolabs), followed by transfection into DH10B cells according to the manufacturer's protocol (GIBCO BRL Products). Once the cDNA inserts are in plasmid vectors, plasmid DNAs are prepared from randomly picked bacterial colonies containing recombinant pBluescript plasmids, or the insert cDNA sequences are amplified via polymerase chain reaction using primers specific for vector sequences flanking the inserted cDNA sequences. Amplified insert DNAs or plasmid DNAs are sequenced in dye-primer sequencing reactions to generate partial cDNA sequences (expressed sequence tags or "ESTs": see Adams et al., (1991) *Science* 252:1651). The resulting ESTs are analyzed using a Perkin Elmer Model 377 fluorescent sequencer.

EXAMPLE 2

Identification of cDNA Clones

cDNA clones encoding 3-dehydroquinase synthase were identified by conducting BLAST (Basic Local Alignment Search Tool; Altschul et al. (1993) *J. Mol. Biol.* 215:403-410; see also www.ncbi.nlm.nih.gov/BLAST/) searches for similarity to sequences contained in the BLAST "nr" database (comprising all non-redundant GenBank CDS translations, sequences derived from the 3-dimensional structure Brookhaven Protein Data Bank, the last major release of the SWISS-PROT protein sequence database, EMBL, and DDBJ databases). The cDNA sequences obtained in Example 1 were analyzed for similarity to all publicly available DNA sequences contained in the "nr" database using the BLASTN algorithm provided by the National Center for Biotechnology Information (NCBI). The DNA sequences were translated in all reading frames and compared for similarity to all publicly available protein sequences contained in the "nr" database using the BLASTX algorithm (Gish and States (1993) *Nature Genetics* 3:266-272) provided by the NCBI. For convenience, the P-value (probability) of observing a match of a cDNA sequence to a sequence contained in the searched databases merely by chance as calculated by BLAST are reported herein as "pLog" values, which represent the negative of the logarithm of the reported P-value. Accordingly, the greater the pLog value, the greater the likelihood that the cDNA sequence and the BLAST "hit" represent homologous proteins.

EXAMPLE 3Characterization of cDNA Clones Encoding 3-Dehydroquinate Synthase

The BLASTX search using the EST sequences from clones listed in Table 3 revealed similarity of the polypeptides encoded by the cDNAs to 3-dehydroquinate synthase from *Escherichia coli* (NCBI General Identifier No. 114181). Shown in Table 3 are the BLAST results for the entire cDNA inserts comprising the indicated cDNA clones ("FIS"), or contigs assembled from an FIS and an EST ("Contig"):

TABLE 3

BLAST Results for Sequences Encoding Polypeptides Homologous to 3-Dehydroquinate Synthase

Clone	Status	BLAST pLog Score 114181
Contig of: cbn2.pk0047.e10 cca.pk0019.e12	Contig	102.00
rls72.pk0035.a10	FIS	102.00
se3.pk0029.f9	FIS	101.00
wr1.pk0011.d5	FIS	79.00

BLASTN analysis of the NCBI EST database indicates that nucleotides 956 to 1300 of the contig assembled from clones cbn2.pk0047.e10 and cca.pk0019.e12 are 99% identical to nucleotides 345 to 1 of a 345 nucleotide corn EST (NCBI General Identifier No. 4688530). The EST was published in NCBI on April 26, 1999. Clone cca.pk0019.e12 encodes the entire 3-dehydroquinate synthase gene but is incompletely processed. Using clone cbn2.pk0047.e10, the sequence corresponding to the intron has been removed. The instant corn 3-dehydroquinate synthase sequence was determined prior to publication of the corn EST in NCBI.

BLASTN analysis of the NCBI EST database indicates that nucleotides 1005 to 1328 from clone rls72.pk0035.a10 are 96% identical to nucleotides 20 to 344 of a 683 nucleotide rice EST (NCBI General Identifier No. 4715434) published on April 29, 1999. Clone rls72.pk0035.a10 encodes an entire 3-dehydroquinate synthase; this sequence was determined prior to publication of the rice EST sequence.

Figure 1 presents an alignment of the amino acid sequences set forth in SEQ ID NOs:2, 4, 6 and 8 and the *Escherichia coli* sequence (SEQ ID NO:9). The data in Table 4 represents a calculation of the percent identity of the amino acid sequences set forth in SEQ ID NOs:2, 4, 6 and 8 and the *Escherichia coli* sequence (SEQ ID NO:9).

TABLE 4

Percent Identity of Amino Acid Sequences Deduced From the Nucleotide Sequences of
cDNA Clones Encoding Polypeptides Homologous
to 3-Dehydroquinate Synthase

SEQ ID NO.	Percent Identity to 114181
2	52.5
4	52.5
6	52.8
8	55.0

Sequence alignments and percent identity calculations were performed using the Megalign program of the LASARGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences was performed using the Clustal method of alignment (Higgins and Sharp (1989) *CABIOS*, 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments using the Clustal method were KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. Sequence alignments and BLAST scores and probabilities indicate that the nucleic acid fragments comprising the instant cDNA clones encode a substantial portion of a 3-dehydroquinate synthase. These sequences represent the first plant sequences encoding 3-dehydroquinate synthase.

EXAMPLE 4Expression of Chimeric Genes in Monocot Cells

A chimeric gene comprising a cDNA encoding the instant polypeptide in sense orientation with respect to the maize 27 kD zein promoter that is located 5' to the cDNA fragment, and the 10 kD zein 3' end that is located 3' to the cDNA fragment, can be constructed. The cDNA fragment of this gene may be generated by polymerase chain reaction (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites (NcoI or SmaI) can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the digested vector pML103 as described below. Amplification is then performed in a standard PCR. The amplified DNA is then digested with restriction enzymes NcoI and SmaI and fractionated on an agarose gel. The appropriate band can be isolated from the gel and combined with a 4.9 kb NcoI-SmaI fragment of the plasmid pML103. Plasmid pML103 has been deposited under the terms of the Budapest Treaty at ATCC (American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209), and bears accession number ATCC 97366. The DNA segment from pML103 contains a 1.05 kb Sall-NcoI promoter fragment of the maize 27 kD zein gene and a 0.96 kb SmaI-Sall fragment from the 3' end of the maize 10 kD zein gene in the vector pGem9Zf(+) (Promega). Vector and insert DNA can be ligated at 15°C overnight,

essentially as described (Maniatis). The ligated DNA may then be used to transform *E. coli* XL1-Blue (Epicurian Coli XL-1 Blue™; Stratagene). Bacterial transformants can be screened by restriction enzyme digestion of plasmid DNA and limited nucleotide sequence analysis using the dideoxy chain termination method (Sequenase™ DNA Sequencing Kit; U.S. Biochemical). The resulting plasmid construct would comprise a chimeric gene encoding, in the 5' to 3' direction, the maize 27 kD zein promoter, a cDNA fragment encoding the instant polypeptide, and the 10 kD zein 3' region.

The chimeric gene described above can then be introduced into corn cells by the following procedure. Immature corn embryos can be dissected from developing caryopses derived from crosses of the inbred corn lines H99 and LH132. The embryos are isolated 10 to 11 days after pollination when they are 1.0 to 1.5 mm long. The embryos are then placed with the axis-side facing down and in contact with agarose-solidified N6 medium (Chu et al. (1975) *Sci. Sin. Peking* 18:659-668). The embryos are kept in the dark at 27°C. Friable embryogenic callus consisting of undifferentiated masses of cells with somatic proembryoids and embryoids borne on suspensor structures proliferates from the scutellum of these immature embryos. The embryogenic callus isolated from the primary explant can be cultured on N6 medium and sub-cultured on this medium every 2 to 3 weeks.

The plasmid, p35S/Ac (obtained from Dr. Peter Eckes, Hoechst Ag, Frankfurt, Germany) may be used in transformation experiments in order to provide for a selectable marker. This plasmid contains the *Pat* gene (see European Patent Publication 0 242 236) which encodes phosphinothricin acetyl transferase (PAT). The enzyme PAT confers resistance to herbicidal glutamine synthetase inhibitors such as phosphinothricin. The *pat* gene in p35S/Ac is under the control of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature* 313:810-812) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*.

The particle bombardment method (Klein et al. (1987) *Nature* 327:70-73) may be used to transfer genes to the callus culture cells. According to this method, gold particles (1 µm in diameter) are coated with DNA using the following technique. Ten µg of plasmid DNAs are added to 50 µL of a suspension of gold particles (60 mg per mL). Calcium chloride (50 µL of a 2.5 M solution) and spermidine free base (20 µL of a 1.0 M solution) are added to the particles. The suspension is vortexed during the addition of these solutions. After 10 minutes, the tubes are briefly centrifuged (5 sec at 15,000 rpm) and the supernatant removed. The particles are resuspended in 200 µL of absolute ethanol, centrifuged again and the supernatant removed. The ethanol rinse is performed again and the particles resuspended in a final volume of 30 µL of ethanol. An aliquot (5 µL) of the DNA-coated gold particles can be placed in the center of a Kapton™ flying disc (Bio-Rad Labs). The particles are then accelerated into the corn tissue with a Biolistic™ PDS-1000/He (Bio-Rad

Instruments, Hercules CA), using a helium pressure of 1000 psi, a gap distance of 0.5 cm and a flying distance of 1.0 cm.

For bombardment, the embryogenic tissue is placed on filter paper over agarose-solidified N6 medium. The tissue is arranged as a thin lawn and covered a circular area of about 5 cm in diameter. The petri dish containing the tissue can be placed in the chamber of the PDS-1000/He approximately 8 cm from the stopping screen. The air in the chamber is then evacuated to a vacuum of 28 inches of Hg. The macrocarrier is accelerated with a helium shock wave using a rupture membrane that bursts when the He pressure in the shock tube reaches 1000 psi.

Seven days after bombardment the tissue can be transferred to N6 medium that contains gluphosinate (2 mg per liter) and lacks casein or proline. The tissue continues to grow slowly on this medium. After an additional 2 weeks the tissue can be transferred to fresh N6 medium containing gluphosinate. After 6 weeks, areas of about 1 cm in diameter of actively growing callus can be identified on some of the plates containing the glufosinate-supplemented medium. These calli may continue to grow when sub-cultured on the selective medium.

Plants can be regenerated from the transgenic callus by first transferring clusters of tissue to N6 medium supplemented with 0.2 mg per liter of 2,4-D. After two weeks the tissue can be transferred to regeneration medium (Fromm et al. (1990) *Bio/Technology* 8:833-839).

EXAMPLE 5

Expression of Chimeric Genes in Dicot Cells

A seed-specific expression cassette composed of the promoter and transcription terminator from the gene encoding the β subunit of the seed storage protein phaseolin from the bean *Phaseolus vulgaris* (Doyle et al. (1986) *J. Biol. Chem.* 261:9228-9238) can be used for expression of the instant polypeptide in transformed soybean. The phaseolin cassette includes about 500 nucleotides upstream (5') from the translation initiation codon and about 1650 nucleotides downstream (3') from the translation stop codon of phaseolin. Between the 5' and 3' regions are the unique restriction endonuclease sites Nco I (which includes the ATG translation initiation codon), Sma I, Kpn I and Xba I. The entire cassette is flanked by Hind III sites.

The cDNA fragment of this gene may be generated by polymerase chain reaction (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the expression vector. Amplification is then performed as described above, and the isolated fragment is inserted into a pUC18 vector carrying the seed expression cassette.

Soybean embryos may then be transformed with the expression vector comprising sequences encoding the instant polypeptide. To induce somatic embryos, cotyledons, 3-5 mm in length dissected from surface sterilized, immature seeds of the soybean cultivar A2872, can be cultured in the light or dark at 26°C on an appropriate agar medium for 6-10 weeks. Somatic embryos which produce secondary embryos are then excised and placed into a suitable liquid medium. After repeated selection for clusters of somatic embryos which multiplied as early, globular staged embryos, the suspensions are maintained as described below.

Soybean embryogenic suspension cultures can be maintained in 35 mL liquid media on a rotary shaker, 150 rpm, at 26°C with fluorescent lights on a 16:8 hour day/night schedule. Cultures are subcultured every two weeks by inoculating approximately 35 mg of tissue into 35 mL of liquid medium.

Soybean embryogenic suspension cultures may then be transformed by the method of particle gun bombardment (Klein et al. (1987) *Nature* (London) 327:70, U.S. Patent No. 4,945,050). A DuPont Biolistic™ PDS1000/HE instrument (helium retrofit) can be used for these transformations.

A selectable marker gene which can be used to facilitate soybean transformation is a chimeric gene composed of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature* 313:810-812), the hygromycin phosphotransferase gene from plasmid pJR225 (from *E. coli*; Gritz et al. (1983) *Gene* 25:179-188) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*. The seed expression cassette comprising the phaseolin 5' region, the fragment encoding the instant polypeptide and the phaseolin 3' region can be isolated as a restriction fragment. This fragment can then be inserted into a unique restriction site of the vector carrying the marker gene.

To 50 µL of a 60 mg/mL 1 µm gold particle suspension is added (in order): 5 µL DNA (1 µg/µL), 20 µL spermidine (0.1 M), and 50 µL CaCl₂ (2.5 M). The particle preparation is then agitated for three minutes, spun in a microfuge for 10 seconds and the supernatant removed. The DNA-coated particles are then washed once in 400 µL 70% ethanol and resuspended in 40 µL of anhydrous ethanol. The DNA/particle suspension can be sonicated three times for one second each. Five µL of the DNA-coated gold particles are then loaded on each macro carrier disk.

Approximately 300-400 mg of a two-week-old suspension culture is placed in an empty 60x15 mm petri dish and the residual liquid removed from the tissue with a pipette. For each transformation experiment, approximately 5-10 plates of tissue are normally bombarded. Membrane rupture pressure is set at 1100 psi and the chamber is evacuated to a vacuum of 28 inches mercury. The tissue is placed approximately 3.5 inches away from the retaining screen and bombarded three times. Following bombardment, the tissue can be divided in half and placed back into liquid and cultured as described above.

Five to seven days post bombardment, the liquid media may be exchanged with fresh media, and eleven to twelve days post bombardment with fresh media containing 50 mg/mL hygromycin. This selective media can be refreshed weekly. Seven to eight weeks post bombardment, green, transformed tissue may be observed growing from untransformed, necrotic embryogenic clusters. Isolated green tissue is removed and inoculated into individual flasks to generate new, clonally propagated, transformed embryogenic suspension cultures. Each new line may be treated as an independent transformation event. These suspensions can then be subcultured and maintained as clusters of immature embryos or regenerated into whole plants by maturation and germination of individual somatic embryos.

EXAMPLE 6

Expression of Chimeric Genes in Microbial Cells

The cDNAs encoding the instant polypeptide can be inserted into the T7 *E. coli* expression vector pBT430. This vector is a derivative of pET-3a (Rosenberg et al. (1987) *Gene* 56:125-135) which employs the bacteriophage T7 RNA polymerase/T7 promoter system. Plasmid pBT430 was constructed by first destroying the EcoR I and Hind III sites in pET-3a at their original positions. An oligonucleotide adaptor containing EcoR I and Hind III sites was inserted at the BamH I site of pET-3a. This created pET-3aM with additional unique cloning sites for insertion of genes into the expression vector. Then, the Nde I site at the position of translation initiation was converted to an Nco I site using oligonucleotide-directed mutagenesis. The DNA sequence of pET-3aM in this region, 5'-CATATGG, was converted to 5'-CCCATGG in pBT430.

Plasmid DNA containing a cDNA may be appropriately digested to release a nucleic acid fragment encoding the protein. This fragment may then be purified on a 1% NuSieve GTG™ low melting agarose gel (FMC). Buffer and agarose contain 10 µg/ml ethidium bromide for visualization of the DNA fragment. The fragment can then be purified from the agarose gel by digestion with GELase™ (Epicentre Technologies) according to the manufacturer's instructions, ethanol precipitated, dried and resuspended in 20 µL of water. Appropriate oligonucleotide adapters may be ligated to the fragment using T4 DNA ligase (New England Biolabs, Beverly, MA). The fragment containing the ligated adapters can be purified from the excess adapters using low melting agarose as described above. The vector pBT430 is digested, dephosphorylated with alkaline phosphatase (NEB) and deproteinized with phenol/chloroform as described above. The prepared vector pBT430 and fragment can then be ligated at 16°C for 15 hours followed by transformation into DH5 electrocompetent cells (GIBCO BRL). Transformants can be selected on agar plates containing LB media and 100 µg/mL ampicillin. Transformants containing the gene encoding the instant polypeptide are then screened for the correct orientation with respect to the T7 promoter by restriction enzyme analysis.

For high level expression, a plasmid clone with the cDNA insert in the correct orientation relative to the T7 promoter can be transformed into *E. coli* strain BL21(DE3) (Studier et al. (1986) *J. Mol. Biol.* 189:113-130). Cultures are grown in LB medium containing ampicillin (100 mg/L) at 25°C. At an optical density at 600 nm of approximately 1, IPTG (isopropylthio- β -galactoside, the inducer) can be added to a final concentration of 0.4 mM and incubation can be continued for 3 h at 25°. Cells are then harvested by centrifugation and re-suspended in 50 μ L of 50 mM Tris-HCl at pH 8.0 containing 0.1 mM DTT and 0.2 mM phenyl methylsulfonyl fluoride. A small amount of 1 mm glass beads can be added and the mixture sonicated 3 times for about 5 seconds each time with a microprobe sonicator. The mixture is centrifuged and the protein concentration of the supernatant determined. One μ g of protein from the soluble fraction of the culture can be separated by SDS-polyacrylamide gel electrophoresis. Gels can be observed for protein bands migrating at the expected molecular weight.

EXAMPLE 7

Evaluating Compounds for Their Ability to Inhibit the Activity of 3-Dehydroquinase Synthase

The polypeptide described herein may be produced using any number of methods known to those skilled in the art. Such methods include, but are not limited to, expression in bacteria as described in Example 6, or expression in eukaryotic cell culture, *in planta*, and using viral expression systems in suitably infected organisms or cell lines. The instant polypeptide may be expressed either as mature forms of the proteins as observed *in vivo* or as fusion proteins by covalent attachment to a variety of enzymes, proteins or affinity tags. Common fusion protein partners include glutathione S-transferase ("GST"), thioredoxin ("Trx"), maltose binding protein, and C- and/or N-terminal hexahistidine polypeptide ("His₆"). The fusion proteins may be engineered with a protease recognition site at the fusion point so that fusion partners can be separated by protease digestion to yield intact mature enzyme. Examples of such proteases include thrombin, enterokinase and factor Xa. However, any protease can be used which specifically cleaves the peptide connecting the fusion protein and the enzyme.

Purification of the instant polypeptide, if desired, may utilize any number of separation technologies familiar to those skilled in the art of protein purification. Examples of such methods include, but are not limited to, homogenization, filtration, centrifugation, heat denaturation, ammonium sulfate precipitation, desalting, pH precipitation, ion exchange chromatography, hydrophobic interaction chromatography and affinity chromatography, wherein the affinity ligand represents a substrate, substrate analog or inhibitor. When the instant polypeptide are expressed as fusion proteins, the purification protocol may include the use of an affinity resin which is specific for the fusion protein tag attached to the expressed enzyme or an affinity resin containing ligands which are specific for the enzyme.

For example, the instant polypeptide may be expressed as a fusion protein coupled to the C-terminus of thioredoxin. In addition, a (His)₆ peptide may be engineered into the N-terminus of the fused thioredoxin moiety to afford additional opportunities for affinity purification. Other suitable affinity resins could be synthesized by linking the appropriate ligands to any suitable resin such as Sepharose-4B. In an alternate embodiment, a thioredoxin fusion protein may be eluted using dithiothreitol; however, elution may be accomplished using other reagents which interact to displace the thioredoxin from the resin. These reagents include β -mercaptoethanol or other reduced thiol. The eluted fusion protein may be subjected to further purification by traditional means as stated above, if desired.

Proteolytic cleavage of the thioredoxin fusion protein and the enzyme may be accomplished after the fusion protein is purified or while the protein is still bound to the ThioBond™ affinity resin or other resin.

Crude, partially purified or purified enzyme, either alone or as a fusion protein, may be utilized in assays for the evaluation of compounds for their ability to inhibit enzymatic activation of the instant polypeptide disclosed herein. Assays may be conducted under well known experimental conditions which permit optimal enzymatic activity. Assays for 3-dehydroquinase synthase are presented by Mehdi et al. (1987) *Methods Enzymol.* 142:306-314; Bender et al. (1989) *Biochemistry* 28:7555-7560 and Gollub et al. (1971) *Methods Enzymol.* 17A:349.

CLAIMS

What is claimed is:

1. An isolated nucleic acid fragment encoding a 3-dehydroquinate synthase comprising a member selected from the group consisting of:

- 5 (a) an isolated nucleic acid fragment encoding an amino acid sequence that is at least 80% identical to the amino acid sequence set forth in a member selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 and SEQ ID NO:8;
- (b) an isolated nucleic acid fragment that is complementary to (a).

10 2. The isolated nucleic acid fragment of Claim 1 wherein nucleic acid fragment is a functional RNA.

3. The isolated nucleic acid fragment of Claim 1 wherein the nucleotide sequence of the fragment comprises the sequence set forth in a member selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 and SEQ ID NO:7.

15 4. A chimeric gene comprising the nucleic acid fragment of Claim 1 operably linked to suitable regulatory sequences.

5. A transformed host cell comprising the chimeric gene of Claim 4.

20 6. A 3-dehydroquinate synthase polypeptide comprising all or a substantial portion of the amino acid sequence set forth in a member selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 and SEQ ID NO:8.

7. A method of altering the level of expression of a 3-dehydroquinate synthase in a host cell comprising:

- 25 (a) transforming a host cell with the chimeric gene of Claim 4 and
- (b) growing the transformed host cell produced in step (a) under conditions that are suitable for expression of the chimeric gene
- wherein expression of the chimeric gene results in production of altered levels of a 3-dehydroquinate synthase in the transformed host cell.

8. A method of obtaining a nucleic acid fragment encoding all or a substantial portion of the amino acid sequence encoding a 3-dehydroquinate synthase comprising:

- 30 (a) probing a cDNA or genomic library with the nucleic acid fragment of Claim 1;
- (b) identifying a DNA clone that hybridizes with the nucleic acid fragment of Claim 1;
- (c) isolating the DNA clone identified in step (b); and
- 35 (d) sequencing the cDNA or genomic fragment that comprises the clone isolated in step (c)

wherein the sequenced nucleic acid fragment encodes all or a substantial portion of the amino acid sequence encoding a 3-dehydroquinate synthase.

9. A method of obtaining a nucleic acid fragment encoding a substantial portion of an amino acid sequence encoding a 3-dehydroquinate synthase comprising:

- (a) synthesizing an oligonucleotide primer corresponding to a portion of the sequence set forth in any of SEQ ID NOs:1, 3, 5 and 7; and
- (b) amplifying a cDNA insert present in a cloning vector using the oligonucleotide primer of step (a) and a primer representing sequences of the cloning vector

wherein the amplified nucleic acid fragment encodes a substantial portion of an amino acid sequence encoding a 3-dehydroquinate synthase.

10. The product of the method of Claim 8.

11. The product of the method of Claim 9.

12. A method for evaluating at least one compound for its ability to inhibit the activity of a 3-dehydroquinate synthase, the method comprising the steps of:

- (a) transforming a host cell with a chimeric gene comprising a nucleic acid fragment encoding a 3-dehydroquinate synthase, operably linked to suitable regulatory sequences;
- (b) growing the transformed host cell under conditions that are suitable for expression of the chimeric gene wherein expression of the chimeric gene results in production of the 3-dehydroquinate synthase encoded by the operably linked nucleic acid fragment in the transformed host cell;
- (c) optionally purifying the 3-dehydroquinate synthase expressed by the transformed host cell;
- (d) treating the 3-dehydroquinate synthase with a compound to be tested; and
- (e) comparing the activity of the 3-dehydroquinate synthase that has been treated with a test compound to the activity of an untreated 3-dehydroquinate synthase,

thereby selecting compounds with potential for inhibitory activity.

SEQ	ID	NO:	9	ME-----	MAASASSLLAAPASSSCGAISPQLPRGAPA----	AASVASPSRHS CYL-LRASPSRRHRS
SEQ	ID	NO:	2			
SEQ	ID	NO:	4		MAAAASSLLAAASSSRAAAVSARRAPSASPA AAAASLPSPSRASCAPPLRASAAATLRS	
SEQ	ID	NO:	6		MASTATNFSLSL---CANQQTPIPKPSFFSNNNNHLHFNSNNNWA----	WASVSTSRKS
SEQ	ID	NO:	8			
				1		60

SEQ	ID NO: 9	RIV-----VTLGRSYPIITASGLFNEPASFLPLKSGEQVML	***	*****	++	++	+++++	+++++	+++++
SEQ	ID NO: 2	RFVANAAPTQMPPAESR--VSTVVDVLDLGRSYPIYGAGLLDEPDLLQQRHVHGRKRLV							
SEQ	ID NO: 4	RVVASAAPAMQPPASR--VSTVVDVLDLGRSYPIYGAGLLDEPDLLQQRHVHGRKRLV							
SEQ	ID NO: 6	RICATSSQVMDPSAAKSEPALPTIVEVDLGRSYPIYGSGLLNQPDYLRHVHGRKRLV							
SEQ	ID NO: 8	-----							
		61							120

```

***      ***      ***      ***      ***      ***      ***      ***      ***      ***      ***
ID NO: 9  VTNETLAPLYLDKVRGVLEQAG--VNVDSVILPDGEQYKSLAVLDTVFTALLQKPHGRDRT
ID NO: 2  VTNTTVAPLYLDKVTWALTNNLNVSVESVILPDGEKYKNMDTLMKFVDKAVESRFDRRRC
ID NO: 4  VTNTTVAPLYLEKVTWALTNNPNVSVESVILPDGEKYKDMGTLMKFVDKAVESRLDRRC
ID NO: 6  VTNETVAPLYLDKVVDDALTRGNPNVSVESVILPDGEQYKMDTLMKFVDKAIESRLDRRC
ID NO: 8  -----AR-----180
121

```

09/744099

FIGURE 1 (CONTINUED)

```

+*****+*****+*****+*****+*****+*****+*****+*****+*****+
SEQ ID NO: 9 TLVALGGVVGDLTGFAAASYQRGVRFIQVPTLLSQVDSSVGCKTAVNHPLGKNMIGAF
SEQ ID NO: 2 TFVALGGV LGDMCGFAAAFLRGVNF IQIPTTLMAQVDSSVGCKTCINHP L GKNLIGAF
SEQ ID NO: 4 TFVALGGVIGDMCGFAAAFLRGVNF IQIPTTLMAQVDSSVGCKTGINHP L GKNLIGAF
SEQ ID NO: 6 TFVALGGVIGDMCGFAAASFLRGVNF IQIPTTVMAQVDSSVGCKTGINHRLGKNMIGTF
SEQ ID NO: 8 AFVALGGVIGDMCGFAAAFLRGVNF IQIPTTLMAQVDSSVGCKTGINHPLGKNLIGAF
181 240

```

SEQ ID NO: 9	* **++**++* +**++**++* **++**++* **++**++* *	YQPASVVVDLCLKTLPPRELASGLAEVIKYGII LDGAFFNWLEENLDALLRLDGPAMAY
SEQ ID NO: 2	YQPQCVLIDTNTLTPDRELASGLAEVVKYGLIRDAPFFEWQEKMPKLLAREPNALAY	
SEQ ID NO: 4	YHPQCVLIDTETLNTLPDRELASGLAEVVKYGLIRDAPFFEWQEKMPALLAREPSALAY	
SEQ ID NO: 6	YQPQCVLIDTDTLNTLPDRELASGLAEVIKYGILIRDAEFFEWQEKMHLLLRDPSVMAY	
SEQ ID NO: 8	YQPQCVLIDTETLNTLPDRELASGLAEVVKYGLIRDAPFFEWQEKMAAILAREPSALTY	
	241	300

	+ + + + + + + + + + * +
SEQ ID NO: 9	CIRRCCELKAEVVAADERETGLRALLNLGHTFGHAI EAEMGYGNWLHG EAVAAAGMVMAAR
SEQ ID NO: 2	AIKRSCENKAEVVAQDEKESGLRATLNLGHTFGHAI ETGTGYGA WLHG EAVAAAGTVMATD
SEQ ID NO: 4	AIKRSCENKAEVVAQDEKESGLRATLNLGHTFGHAI ETGTGYGA WLHG EAVAAAGTVM AAD
SEQ ID NO: 6	AIKRSCENKAEVVSLDQESGLRATLNLGHTFGHAI ETGVGYGQW LHGEA VAAAGTVMAVD
SEQ ID NO: 8	AIKRSCENKAEVVAQDEKESGLRATLNLGHTFGHAI ETGLGYGE W LHG EAVAAAGTVM AAD
	301 360

SEQ ID NO: 9	TSERLGQFSSAETQRIITLLKRALPVNGPREMSAQAYLPHMLRDKKVLAGEMLRIL-PL
SEQ ID NO: 2	MSHRLGWIDDSIRKRVVDILKQAKLP IAPPETMTVLIFKNIMAVDKKVADGLRLILLKG
SEQ ID NO: 4	MSHRLGWIDESIKKRAIDILEKAKLPITPPEAMTVEKEKFSIMAVDKKVADGLRLILLKG
SEQ ID NO: 6	MSYRLGWIDDSLVKRVGDILKQAKLP TAPPETVTVDMFKSVMAVDKKVADGLRLILLKG
SEQ ID NO: 8	MSYRLGWIDESIKKRTFDILDQAKLPVTS PKGMTVEKEFERNIMAVDKKVADGLRLILLKG

3/3

PTO/PGT Rec'd 27 JUN 2001

#15

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN THE APPLICATION OF:
R. CAHOON ET AL.

CASE NO.: BB1159 US PCT2

APPLICATION NO.: 09/744099

GROUP ART UNIT: UNKNOWN

FILED: JANUARY 16, 2001

EXAMINER: UNKNOWN

FOR: **CHORISMATE BIOSYNTHESIS ENZYMES**

Assistant Commissioner for Patents
Washington, DC 20231

Sir:

DECLARATION IN ACCORDANCE WITH 37 CFR 1.821

I hereby state that the content of the paper and computer readable copies of the Sequence Listing, submitted in accordance with 37 CFR 1.821(c) and (e), respectively are the same.

Respectfully submitted,



KENNING LI
ATTORNEY FOR APPLICANTS
REGISTRATION NO. 44,872
TELEPHONE: 302-992-3749
FACSIMILE: 302-892-1026

Dated: 06/25/2001

09744099-011601

DECLARATION and POWER OF ATTORNEY

As a below-named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

CHORISMATE BIOSYNTHESIS ENZYMES

the specification of which is attached hereto unless the following box is checked:

☒ was filed on **20 JULY 1999** as U.S. Application No. _____ or PCT International Application No. _____

PCT/US99/16354 and was amended on _____ (if applicable).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is known to me to be material to patentability as defined in 37 CFR § 1.56.

I hereby claim foreign priority benefits under 35 U.S.C. § 119(a)-(d) or § 365(b) of any foreign application(s) for patent or inventor's certificate, or § 365(a) of any PCT International application which designated at least one country other than the United States, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate, or PCT International application having a filing date before that of the application on which priority is claimed.

Application No.	Country	Filing Date	Priority Claimed (Yes/No)
-----------------	---------	-------------	---------------------------

I hereby claim the benefit under 35 U.S.C. § 119(e) of any United States Provisional Application(s) listed below.

U.S. Provisional Application No.

60/093611

U.S. Filing Date

21 JULY 1998

I hereby claim the benefit under 35 U.S.C. § 120 of any United States application(s), or § 365(c) of any PCT International Application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application or PCT International Application in the manner provided by the first paragraph of 35 U.S.C. § 112, I acknowledge the duty to disclose information which is known to me to be material to patentability as defined in 37 CFR § 1.56 which became available between the filing date of the prior application and the national or PCT International filing date of this application.

Application No.	Filing Date	Status (patented, pending or abandoned)
-----------------	-------------	---

POWER OF ATTORNEY: I hereby appoint the following attorney(s) and/or agent(s) the power to prosecute this application and transact all business in the Patent and Trademark Office connected therewith:

Name: **WILLIAM R. MAJARIAN**

Registration No.: **41,173**

Send correspondence and direct telephone calls to:

WILLIAM R. MAJARIAN

E. I. du Pont de Nemours and Company
Legal - Patents
Wilmington, DE 19898, U.S.A.

Tel. No.
(302) 992-4926
Fax No.
(302) 892-7949

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

INVENTOR(S)

Full Name of Inventor	Last Name CAHOON	First Name REBECCA	Middle Name E.
	Signature (please sign full name) <i>Rebecca E. Cahoon</i>		Date <i>August 26, 1999</i>
Residence & Citizenship	City WILMINGTON	State or Foreign Country DELAWARE	Country of Citizenship U.S.A.
Post Office Address	Post Office Address 2331 WEST 18TH STREET	City WILMINGTON	State or Country DELAWARE
			Zip Code 19806
Full Name of Inventor	Last Name KINNEY	First Name ANTHONY	Middle Name JOHN
	Signature (please sign full name) <i>Anthony J. Kinney</i>		Date <i>8/30/99</i>
Residence & Citizenship	City WILMINGTON	State or Foreign Country DELAWARE	Country of Citizenship GB
Post Office Address	Post Office Address 609 LORE AVENUE	City WILMINGTON	State or Country DELAWARE
			Zip Code 19809
Full Name of Inventor	Last Name RAFALSKI	First Name J.	Middle Name ANTONI
	Signature (please sign full name) <i>J. Antoni Rafalski</i>		Date <i>8/31/99</i>
Residence & Citizenship	City WILMINGTON	State or Foreign Country DELAWARE	Country of Citizenship U.S.A.
Post Office Address	Post Office Address 2028 LONGCOME DRIVE	City WILMINGTON	State or Country DELAWARE
			Zip Code 19810

Full Name of Inventor <i>HP</i>	Last Name REDINA <i>Rendina</i>	First Name ALAN	Middle Name R.
	Signature (please sign full name): <i>Alan R. Rendina</i>		Date: <i>9/8/99</i>
Residence & Citizenship	City WILMINGTON <i>DE</i>	State or Foreign Country DELAWARE	Country of Citizenship U.S.A.
Post Office Address	Post Office Address 506 RIBLETT LANE	City WILMINGTON	State or Country DELAWARE Zip Code 19808

09744099.011601
T09T01504460

SEQUENCE LISTING

<110> E. I. du Pont de Nemours and Company

<120> Chorismate Biosynthesis Enzymes

<130> BB-1159

<140>

<141>

<150> 60/093,611

<151> July 21, 1998

<160> 9

<170> Microsoft Office 97

<210> 1

<211> 1628

<212> DNA

<213> Zea mays

<400> 1

```

ccgccgcttt cccagctctc atggcggcgt ccgcgtctc cctccttgcc gcccccgcat 60
cctcatcctg tggggccatc tccccccagc tccctcgtgg agctccggcc gctgcctcgg 120
tcgcttcgcc gtcgcgtcac tcctgctaoc tcctccgcgc tcccccttcg aggcgccatc 180
ggagtcgctt cgtagccaac gccgctccca ccatgcagcc tccggcggag tccagggtct 240
ccacggtagt cgatgtcgac ctgcggcgacc gtagctatcc gatctacac gccgcaggcc 300
tcctcgacga gccggacctg ctgcagaggc atgttcatgg taagagggtt ctggtggtga 360
ccaacacgac cgtcgcgcgg ctttacctgg acaagggtgac atgggcactc acccacaaca 420
acctgaatgt atcagtgga agcgtgatcc tgcccgcagg tgaaggtac aaaaatatgg 480
acacgctgat gaagggtgtt gacaaggcag tcgagtcccg ttttgaccgc cgggtgcacat 540
ttgtagcact ggggtggtgg gtcatgggg acatgtgtgg ttttgacgct gctgcattcc 600
tcgggggcgt caatttcata cagataccaa ctactctgat ggcccagggt gattcatctg 660
ttggtgggaa aaccgggatt aaccacccac taggaaagaa cttgattggg gcattctacc 720
agccacaatg tgttctaatt gacacaaata cacttaacac attgcctgac agggagctag 780
cttcaggcat tgccgaggta gtaaagtatg ggctcataag ggatgcacca ttctttgagt 840
ggcaagagaa gaacatgccg aaattgttag caagagaacc aaatgctttg gcatatgcta 900
tcaagagatc atgtgaaaac aaagctgaag tgggtggcaca agatgagaag gaaagtggcc 960
ttcgagcaac actaaacctg ggtcacacat ttggccatgc tattgagact gggacaggat 1020
atggagcatg gctccatggg gaggtgtcg cagctggaac agttatggca actgacatgt 1080
ctcaccgcct ggggtggata gatgactcca tcagaaaacg tgtggttgac atactaaagc 1140
aagccaaact tcccattgca cctcctgaga ccatgaccgt agagaagttt aaaaacatca 1200
tggctgttga caagaagggt gctgatggtc tgttgagact catccttctg aaaggaccgc 1260
taggtgtctg tgtatttacg ggggattatg acgggaatgc actcgatgaa accctacatg 1320
cattctgcga caactgagac atcccagttt tggacatcac ttctgtatgc tagtactagg 1380
ctatgtatca atcagagaaa atattgtggt gtacattatc ttcactgcct catttgtata 1440
ttttttgtca gaaataatgg tatattgttt taagtgtctg tttgagtttg tagagaacca 1500
tggctaataca actgcggtea gttgttctct agtaataata gtaagaagc ttcactacat 1560
cccttctgtc tgttactgtc aagtaacaat actcgtagag atttaaacca aaaaaaaaaa 1620
aaaaaaaaa

```

<210> 2

<211> 438

<212> PRT

<213> Zea mays

<400> 2

```

Met Ala Ala Ser Ala Ser Ser Leu Leu Ala Ala Pro Ala Ser Ser Ser
 1          5          10          15

Cys Gly Ala Ile Ser Pro Gln Leu Pro Arg Gly Ala Pro Ala Ala Ala
 20          25          30

Ser Val Ala Ser Pro Ser Arg His Ser Cys Tyr Leu Leu Arg Ala Ser
 35          40          45

Pro Ser Arg Arg His Arg Ser Arg Phe Val Ala Asn Ala Ala Pro Thr
 50          55          60

Met Gln Pro Pro Ala Glu Ser Arg Val Ser Thr Val Val Asp Val Asp
 65          70          75          80

Leu Gly Asp Arg Ser Tyr Pro Ile Tyr Ile Gly Ala Gly Leu Leu Asp
 85          90          95

Glu Pro Asp Leu Leu Gln Arg His Val His Gly Lys Arg Val Leu Val
100          105          110

Val Thr Asn Thr Thr Val Ala Pro Leu Tyr Leu Asp Lys Val Thr Trp
115          120          125

Ala Leu Thr His Asn Asn Leu Asn Val Ser Val Glu Ser Val Ile Leu
130          135          140

Pro Asp Gly Glu Lys Tyr Lys Asn Met Asp Thr Leu Met Lys Val Phe
145          150          155          160

Asp Lys Ala Val Glu Ser Arg Phe Asp Arg Arg Cys Thr Phe Val Ala
165          170          175

Leu Gly Gly Gly Val Ile Gly Asp Met Cys Gly Phe Ala Ala Ala Ala
180          185          190

Phe Leu Arg Gly Val Asn Phe Ile Gln Ile Pro Thr Thr Leu Met Ala
195          200          205

Gln Val Asp Ser Ser Val Gly Gly Lys Thr Gly Ile Asn His Pro Leu
210          215          220

Gly Lys Asn Leu Ile Gly Ala Phe Tyr Gln Pro Gln Cys Val Leu Ile
225          230          235          240

Asp Thr Asn Thr Leu Asn Thr Leu Pro Asp Arg Glu Leu Ala Ser Gly
245          250          255

Ile Ala Glu Val Val Lys Tyr Gly Leu Ile Arg Asp Ala Pro Phe Phe
260          265          270

Glu Trp Gln Glu Lys Asn Met Pro Lys Leu Leu Ala Arg Glu Pro Asn
275          280          285

Ala Leu Ala Tyr Ala Ile Lys Arg Ser Cys Glu Asn Lys Ala Glu Val
290          295          300

Val Ala Gln Asp Glu Lys Glu Ser Gly Leu Arg Ala Thr Leu Asn Leu
305          310          315          320

```

Gly His Thr Phe Gly His Ala Ile Glu Thr Gly Thr Gly Tyr Gly Ala
325 330 335

Trp Leu His Gly Glu Ala Val Ala Ala Gly Thr Val Met Ala Thr Asp
340 345 350

Met Ser His Arg Leu Gly Trp Ile Asp Asp Ser Ile Arg Lys Arg Val
355 360 365

Val Asp Ile Leu Lys Gln Ala Lys Leu Pro Ile Ala Pro Pro Glu Thr
370 375 380

Met Thr Val Glu Lys Phe Lys Asn Ile Met Ala Val Asp Lys Lys Val
385 390 395 400

Ala Asp Gly Leu Leu Arg Leu Ile Leu Leu Lys Gly Pro Leu Gly Cys
405 410 415

Cys Val Phe Thr Gly Asp Tyr Asp Gly Asn Ala Leu Asp Glu Thr Leu
420 425 430

His Ala Phe Cys Asp Asn
435

<210> 3
<211> 1358
<212> DNA
<213> Oryza sativa

<400> 3
gcacgagtac agcgtcgtct tcgttttcgca gctccaccac ctccatggcg gccgcgcgct 60
cctcctctct gtcgcgccgc gccctcgtct cctcccgcgc gccgcccgtg tccgcccgc 120
gcgccccacg cgcgagtcgg gccgcggcgg cctcccctcc ctgcgcgtct cgcgcctcct 180
gcgtcctctc cctccgcgct tccgcagcga ggaccctccg cagccgtgtc gtcgcgagcg 240
cgccccccgc tatgcagccg ccgcggcggt cgagggtctc cagcgttgtc gacgtcgacc 300
tcggcgaccg gagctaccgg atctacatcg gcgcaggcct tctcgacgag cctgacctgc 360
tgcagaggca tgttcatggg aagagggttt tgggtggtgac caacaccacc gtcgcgccgc 420
tctacctgga gaagggtgacc tgggcactca cgcacaacaa cccgaatgtt tctgtggaga 480
gcgtgatcct gcccgacggc gagaagtaca aggacatggg cacactgatg aagggttttcg 540
acaaggcagt cgagtcccgg ctggaccggc ggtgcacgtt tgttgcggtg ggaggtggcg 600
ttattgggga catgtgcggt tttgcagctg ctgcgttctt gcgtggtgtc aatttcatac 660
agattcctac tactctgatg gcccggtgg attcatctgt tggagggaag accggcatta 720
accatccatt ggggaagaac ttaattgggg cgttctacca cccacagtgt gtactgatag 780
acaccgagac actgaatata ttgcctgaca gggaactggc ttcaggcata gctgaggtgg 840
tgaagtatgg tctcataaga gatgcaccgt tctttgaatg gcaagagaaa aacatgccag 900
cattatttagc aagagaacca agtgctctgg cctatgctat taagagatcg tgtgaaaaca 960
aagctgaagt ggttgctcag gacgagaagg aaagtgggtc ccgagcaaca cttaatctcg 1020
gccatacatt tggccatgct atagaaacag gaactggcta tggagcatgg ctccatgggg 1080
aggctgttgc agctggaaca gttatggcag ctgacatgtc tcaccgcctg ggttggatag 1140
acgagtcaat caagaaacgg gcaattgaca tactagagaa agcgaagctt ccaattacac 1200
ctccagaggc catgacagtg gagaagttca aaagtattat ggccgttgat aagaaggttg 1260
ctgatggatt gctgaggctc atcctcctga aaggacctct gggaagctgt gttttcactg 1320
gcgattactg ttcttcgagg tcgacgtgtc gataaatt 1358

<210> 4
<211> 436
<212> PRT
<213> Oryza sativa

<400> 4
 Met Ala Ala Ala Ala Ser Ser Ser Leu Leu Ala Ala Ala Ser Ser Ser
 1 5 10 15
 Ser Arg Ala Ala Ala Val Ser Ala Arg Arg Ala Pro Ser Ala Ser Pro
 20 25 30
 Ala Ala Ala Ala Ser Leu Pro Ser Pro Ser Arg Ala Ser Cys Ala Pro
 35 40 45
 Pro Leu Arg Ala Ser Ala Ala Arg Thr Leu Arg Ser Arg Val Val Ala
 50 55 60
 Ser Ala Ala Pro Ala Met Gln Pro Pro Pro Ala Ser Arg Val Ser Thr
 65 70 75 80
 Val Val Asp Val Asp Leu Gly Asp Arg Ser Tyr Pro Ile Tyr Ile Gly
 85 90 95
 Ala Gly Leu Leu Asp Glu Pro Asp Leu Leu Gln Arg His Val His Gly
 100 105 110
 Lys Arg Val Leu Val Val Thr Asn Thr Thr Val Ala Pro Leu Tyr Leu
 115 120 125
 Glu Lys Val Thr Trp Ala Leu Thr His Asn Asn Pro Asn Val Ser Val
 130 135 140
 Glu Ser Val Ile Leu Pro Asp Gly Glu Lys Tyr Lys Asp Met Gly Thr
 145 150 155 160
 Leu Met Lys Val Phe Asp Lys Ala Val Glu Ser Arg Leu Asp Arg Arg
 165 170 175
 Cys Thr Phe Val Ala Leu Gly Gly Gly Val Ile Gly Asp Met Cys Gly
 180 185 190
 Phe Ala Ala Ala Ala Phe Leu Arg Gly Val Asn Phe Ile Gln Ile Pro
 195 200 205
 Thr Thr Leu Met Ala Gln Val Asp Ser Ser Val Gly Gly Lys Thr Gly
 210 215 220
 Ile Asn His Pro Leu Gly Lys Asn Leu Ile Gly Ala Phe Tyr His Pro
 225 230 235 240
 Gln Cys Val Leu Ile Asp Thr Glu Thr Leu Asn Thr Leu Pro Asp Arg
 245 250 255
 Glu Leu Ala Ser Gly Ile Ala Glu Val Val Lys Tyr Gly Leu Ile Arg
 260 265 270
 Asp Ala Pro Phe Phe Glu Trp Gln Glu Lys Asn Met Pro Ala Leu Leu
 275 280 285
 Ala Arg Glu Pro Ser Ala Leu Ala Tyr Ala Ile Lys Arg Ser Cys Glu
 290 295 300
 Asn Lys Ala Glu Val Val Ala Gln Asp Glu Lys Glu Ser Gly Leu Arg
 305 310 315 320

Ala Thr Leu Asn Leu Gly His Thr Phe Gly His Ala Ile Glu Thr Gly
325 330 335

Thr Gly Tyr Gly Ala Trp Leu His Gly Glu Ala Val Ala Ala Gly Thr
340 345 350

Val Met Ala Ala Asp Met Ser His Arg Leu Gly Trp Ile Asp Glu Ser
355 360 365

Ile Lys Lys Arg Ala Ile Asp Ile Leu Glu Lys Ala Lys Leu Pro Ile
370 375 380

Thr Pro Pro Glu Ala Met Thr Val Glu Lys Phe Lys Ser Ile Met Ala
385 390 395 400

Val Asp Lys Lys Val Ala Asp Gly Leu Leu Arg Leu Ile Leu Leu Lys
405 410 415

Gly Pro Leu Gly Ser Cys Val Phe Thr Gly Asp Tyr Cys Ser Ser Arg
420 425 430

Ser Thr Cys Arg
435

<210> 5

<211> 1643

<212> DNA

<213> Glycine max

<400> 5

```
gcacgagccc ggagtgggtg ttgtgttcaa ggaagacagg cgcaggcaca ggcacatcat 60
gttattgtta ttgttggttt cttcactcag caaccaccaa accacaccgt cgcacccatt 120
ccatcgctga ttcccccaat ggcttccact gccaccaatt tctctctttc tctctgcgcc 180
aaccaacaaa ctccaatccc caaacccctct ttcttctcca ataacaacca tttgcacttc 240
aactctaata ataattgggc ctgggcctct gtttccacct ctgcgaagtc aaggatatgc 300
gccacctcct ctcaagttat ggatccctcc gcagcaaaat ccgaaccgcg tcttcccacc 360
atcgtcgaag tcgatttggt tagccggagc tatcctatct acatcggtat cgggttacta 420
aaccaaccgg actatctcca gaggcattgt catggaaaga gattcctagt tgtaactaac 480
gaaaccgttg cgccacttta tctagacaag ttgtttgatg ctttgacaag gggaaacccg 540
aatgtttctg tggagagtgt aattttacct gatggtgagc agtacaagga catggatatt 600
cttatgaaag tctttgacaa ggccatcgag tcgcgctcgc accggcgggt tacatttggt 660
gctcttggtg gtggtgtgat tggcgacatg tgtggctttg ctgcctctgc cttcctacgt 720
ggtgttaatt ttattcagat tcttacaact gtgatggcac aggtcgattc ttcagttggt 780
ggaaaaactg ggataaatca ccgccttggg aagaacatga tcggtacctt ttaccaacct 840
cagtgtgtgc ttatagacac agacacatta aatacgctac cggataggga actggcatca 900
gggctagcag aggttataaa gtatgggctc attagggatg cagagttttt tgagtggcaa 960
gagaaaaata tgcacttatt attggcaaga gatcctagtg taatggcata tgctataaag 1020
cgatcttgtg aaaacaaggc tgaggttgtg tcttagatc agaaggaaaag tggactgagg 1080
gcaacattga acttgggtca tacatttggt catgcaatag aaactggggg aggctatggg 1140
cagtggcttc atggagaggc tgttgacgct ggcacggtaa tggctgttga catgtcatat 1200
cgcttaggtt ggattgatga ttctcttggt aaacgagttg gagacatttt aaaacaggct 1260
aagttaccca cagccctccc tgagaccgtg actgtggaca tgtttaaatc tgtcatggca 1320
gtggataaga aggtagcaga tgggttgcta aggcctatcc ttctaaaggg tctcttaggc 1380
aattgtgttt tcacagggga ttatgacaga aaggctctag acaatacgct tcgtgcattc 1440
tgtaaatcct gatctttgtg ctccatgttt ctattgcagc tttatccaat tcttttagt 1500
gagtttttgt atgttggtat ataacctcac ttgtattttt cttatgtact ttggatggca 1560
accatgggtc ccattatata cgcaagagta tattatttga aattgaggca tgatcaataa 1620
ttattaaaaa aaaaaaaaaa aaa 1643
```

<210> 6
 <211> 437
 <212> PRT
 <213> Glycine max

<400> 6
 Met Ala Ser Thr Ala Thr Asn Phe Ser Leu Ser Leu Cys Ala Asn Gln
 1 5 10 15
 Gln Thr Pro Ile Pro Lys Pro Ser Phe Phe Ser Asn Asn Asn His Leu
 20 25 30
 His Phe Asn Ser Asn Asn Asn Trp Ala Trp Ala Ser Val Ser Thr Ser
 35 40 45
 Arg Lys Ser Arg Ile Cys Ala Thr Ser Ser Gln Val Met Asp Pro Ser
 50 55 60
 Ala Ala Lys Ser Glu Pro Ala Leu Pro Thr Ile Val Glu Val Asp Leu
 65 70 75 80
 Gly Ser Arg Ser Tyr Pro Ile Tyr Ile Gly Ser Gly Leu Leu Asn Gln
 85 90 95
 Pro Asp Tyr Leu Gln Arg His Val His Gly Lys Arg Val Leu Val Val
 100 105 110
 Thr Asn Glu Thr Val Ala Pro Leu Tyr Leu Asp Lys Val Val Asp Ala
 115 120 125
 Leu Thr Arg Gly Asn Pro Asn Val Ser Val Glu Ser Val Ile Leu Pro
 130 135 140
 Asp Gly Glu Gln Tyr Lys Asp Met Asp Thr Leu Met Lys Val Phe Asp
 145 150 155 160
 Lys Ala Ile Glu Ser Arg Leu Asp Arg Arg Cys Thr Phe Val Ala Leu
 165 170 175
 Gly Gly Gly Val Ile Gly Asp Met Cys Gly Phe Ala Ala Ser Ala Phe
 180 185 190
 Leu Arg Gly Val Asn Phe Ile Gln Ile Pro Thr Thr Val Met Ala Gln
 195 200 205
 Val Asp Ser Ser Val Gly Gly Lys Thr Gly Ile Asn His Arg Leu Gly
 210 215 220
 Lys Asn Met Ile Gly Thr Phe Tyr Gln Pro Gln Cys Val Leu Ile Asp
 225 230 235 240
 Thr Asp Thr Leu Asn Thr Leu Pro Asp Arg Glu Leu Ala Ser Gly Leu
 245 250 255
 Ala Glu Val Ile Lys Tyr Gly Leu Ile Arg Asp Ala Glu Phe Phe Glu
 260 265 270
 Trp Gln Glu Lys Asn Met His Leu Leu Leu Ala Arg Asp Pro Ser Val
 275 280 285

Met Ala Tyr Ala Ile Lys Arg Ser Cys Glu Asn Lys Ala Glu Val Val
 290 295 300

Ser Leu Asp Gln Lys Glu Ser Gly Leu Arg Ala Thr Leu Asn Leu Gly
 305 310 315 320

His Thr Phe Gly His Ala Ile Glu Thr Gly Val Gly Tyr Gly Gln Trp
 325 330 335

Leu His Gly Glu Ala Val Ala Ala Gly Thr Val Met Ala Val Asp Met
 340 345 350

Ser Tyr Arg Leu Gly Trp Ile Asp Asp Ser Leu Val Lys Arg Val Gly
 355 360 365

Asp Ile Leu Lys Gln Ala Lys Leu Pro Thr Ala Pro Pro Glu Thr Val
 370 375 380

Thr Val Asp Met Phe Lys Ser Val Met Ala Val Asp Lys Lys Val Ala
 385 390 395 400

Asp Gly Leu Leu Arg Leu Ile Leu Leu Lys Gly Pro Leu Gly Asn Cys
 405 410 415

Val Phe Thr Gly Asp Tyr Asp Arg Lys Ala Leu Asp Asn Thr Leu Arg
 420 425 430

Ala Phe Cys Lys Ser
 435

<210> 7
 <211> 1103
 <212> DNA
 <213> Triticum aestivum

<400> 7
 gcacgagcat tctgttgcgt aggtggcggt gtcattgggg acatgtgtgg gtttgcagct 60
 gctgcattcc tccgtgggtg caatttcata caaataccca caactctgat ggcccagggtg 120
 gattcttctg ttggagggaa gactgggatt aaccatccac tgggaaagaa ctaattggc 180
 gcattctacc agccgcagtg tgtactcatt gacacagaga cactgaatac attgcctgac 240
 agagagctgg cttcagggtg cgtcgagggt gtgaagtatg gtctcataag agatgcaccc 300
 ttctttgagt ggcaagaaaa gaatatggca gcaatcttag cgagagaacc aagtgccctg 360
 acctacgcta taaagagatc atgtgaaaaa aaagctgaag ttgttgctca ggatgagaag 420
 gaaagtggtc ttcgagcaac actaaacctg ggtcatacat ttggtcatgc tatagaaaca 480
 gggcttggct atggagaatg gctccacggg gaggtgttg ctgctggaac ggttatggca 540
 gctgacatgt cttaccgcct gggctggata gacgagtcca tcaagaaacg gacatttgac 600
 atactagatc aagccaagct tcccgtgaca tcaccaaagg gcatgacggt agagaagttc 660
 agaaacatca tggcgggtcg caagaaggtc gcggacggat tgttgaggct catccttctg 720
 aaaggacctc tgggaggctg cgttttcacc ggcgagtacg acaggaaagc cctggacgaa 780
 accctccgcg ctttctgcga caactgatcg ttgtcaagct ggagctccag catttctgta 840
 tactagtagg ctgtaatatc atatcagaga gagagagaga gagagtgcac tgtgtgctac 900
 ctttattttt tcatttcggt ttggtgagaa taatggaatg cggtaagggg ttggttccag 960
 gtcgtagtac taccaattca ttgtagggca tcatgcctca gtaatcgtga atcttgatca 1020
 gtcattttga gaaacacaat tcggaacgag atcctctgac gtaggaagtt ttaaaaaaag 1080
 agaagaaaaa aaaaaaaaaa aaa 1103

<210> 8
 <211> 268
 <212> PRT
 <213> Triticum aestivum

<400> 8

Ala Arg Ala Phe Val Ala Leu Gly Gly Gly Val Ile Gly Asp Met Cys
 1 5 10 15

Gly Phe Ala Ala Ala Phe Leu Arg Gly Val Asn Phe Ile Gln Ile
 20 25 30

Pro Thr Thr Leu Met Ala Gln Val Asp Ser Ser Val Gly Gly Lys Thr
 35 40 45

Gly Ile Asn His Pro Leu Gly Lys Asn Leu Ile Gly Ala Phe Tyr Gln
 50 55 60

Pro Gln Cys Val Leu Ile Asp Thr Glu Thr Leu Asn Thr Leu Pro Asp
 65 70 75 80

Arg Glu Leu Ala Ser Gly Val Ala Glu Val Val Lys Tyr Gly Leu Ile
 85 90 95

Arg Asp Ala Pro Phe Phe Glu Trp Gln Glu Lys Asn Met Ala Ala Ile
 100 105 110

Leu Ala Arg Glu Pro Ser Ala Leu Thr Tyr Ala Ile Lys Arg Ser Cys
 115 120 125

Glu Asn Lys Ala Glu Val Val Ala Gln Asp Glu Lys Glu Ser Gly Leu
 130 135 140

Arg Ala Thr Leu Asn Leu Gly His Thr Phe Gly His Ala Ile Glu Thr
 145 150 155 160

Gly Leu Gly Tyr Gly Glu Trp Leu His Gly Glu Ala Val Ala Ala Gly
 165 170 175

Thr Val Met Ala Ala Asp Met Ser Tyr Arg Leu Gly Trp Ile Asp Glu
 180 185 190

Ser Ile Lys Lys Arg Thr Phe Asp Ile Leu Asp Gln Ala Lys Leu Pro
 195 200 205

Val Thr Ser Pro Lys Gly Met Thr Val Glu Lys Phe Arg Asn Ile Met
 210 215 220

Ala Val Asp Lys Lys Val Ala Asp Gly Leu Leu Arg Leu Ile Leu Leu
 225 230 235 240

Lys Gly Pro Leu Gly Gly Cys Val Phe Thr Gly Glu Tyr Asp Arg Lys
 245 250 255

Ala Leu Asp Glu Thr Leu Arg Ala Phe Cys Asp Asn
 260 265

<210> 9

<211> 362

<212> PRT

<213> Escherichia coli

<400> 9

```

Met Glu Arg Ile Val Val Thr Leu Gly Glu Arg Ser Tyr Pro Ile Thr
 1             5             10             15

Ile Ala Ser Gly Leu Phe Asn Glu Pro Ala Ser Phe Leu Pro Leu Lys
      20             25             30

Ser Gly Glu Gln Val Met Leu Val Thr Asn Glu Thr Leu Ala Pro Leu
      35             40             45

Tyr Leu Asp Lys Val Arg Gly Val Leu Glu Gln Ala Gly Val Asn Val
      50             55             60

Asp Ser Val Ile Leu Pro Asp Gly Glu Gln Tyr Lys Ser Leu Ala Val
      65             70             75             80

Leu Asp Thr Val Phe Thr Ala Leu Leu Gln Lys Pro His Gly Arg Asp
      85             90             95

Thr Thr Leu Val Ala Leu Gly Gly Gly Val Val Gly Asp Leu Thr Gly
      100            105            110

Phe Ala Ala Ala Ser Tyr Gln Arg Gly Val Arg Phe Ile Gln Val Pro
      115            120            125

Thr Thr Leu Leu Ser Gln Val Asp Ser Ser Val Gly Gly Lys Thr Ala
      130            135            140

Val Asn His Pro Leu Gly Lys Asn Met Ile Gly Ala Phe Tyr Gln Pro
      145            150            155            160

Ala Ser Val Val Val Asp Leu Asp Cys Leu Lys Thr Leu Pro Pro Arg
      165            170            175

Glu Leu Ala Ser Gly Leu Ala Glu Val Ile Lys Tyr Gly Ile Ile Leu
      180            185            190

Asp Gly Ala Phe Phe Asn Trp Leu Glu Glu Asn Leu Asp Ala Leu Leu
      195            200            205

Arg Leu Asp Gly Pro Ala Met Ala Tyr Cys Ile Arg Arg Cys Cys Glu
      210            215            220

Leu Lys Ala Glu Val Val Ala Ala Asp Glu Arg Glu Thr Gly Leu Arg
      225            230            235            240

Ala Leu Leu Asn Leu Gly His Thr Phe Gly His Ala Ile Glu Ala Glu
      245            250            255

Met Gly Tyr Gly Asn Trp Leu His Gly Glu Ala Val Ala Ala Gly Met
      260            265            270

Val Met Ala Ala Arg Thr Ser Glu Arg Leu Gly Gln Phe Ser Ser Ala
      275            280            285

Glu Thr Gln Arg Ile Ile Thr Leu Leu Lys Arg Ala Gly Leu Pro Val
      290            295            300

Asn Gly Pro Arg Glu Met Ser Ala Gln Ala Tyr Leu Pro His Met Leu
      305            310            315            320

```

Arg Asp Lys Lys Val Leu Ala Gly Glu Met Arg Leu Ile Leu Pro Leu
325 330 335

Ala Ile Gly Lys Ser Glu Val Arg Ser Gly Val Ser His Glu Leu Val
340 345 350

Leu Asn Ala Ile Ala Asp Cys Gln Ser Ala
355 360

PCT09

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/744,099

DATE: 08/01/2001
 TIME: 18:19:59

Input Set : A:\BB1159 SEQ LST.txt
 Output Set: N:\CRF3\08012001\I744099.raw

ENTERED

3 <110> APPLICANT: E. I. du Pont de Nemours and Company
 5 <120> TITLE OF INVENTION: Chorismate Biosynthesis Enzymes
 7 <130> FILE REFERENCE: BB-1159

C--> 9 <140> CURRENT APPLICATION NUMBER: US/09/744,099

C--> 10 <141> CURRENT FILING DATE: 2001-06-27

12 <150> PRIOR APPLICATION NUMBER: 60/093,611

W--> 13 <151> PRIOR FILING DATE: July 21, 1998 ←

15 <160> NUMBER OF SEQ ID NOS: 9

17 <170> SOFTWARE: Microsoft Office 97

19 <210> SEQ ID NO: 1

20 <211> LENGTH: 1628

21 <212> TYPE: DNA

22 <213> ORGANISM: Zea mays

24 <400> SEQUENCE: 1

```

25 ccgcgcgttt ccgagctctc atggcggcgt ccgcgtcttc cctccttgcc gcccccgcat 60
26 cctcatcctg tggggccatc tccccccagc tccctcgtgg agctccggcc gctgcctccg 120
27 tcgcttgcgc gtcgcgtcac tcttctacc tctccgcgc tcccccttcg aggcgccatc 180
28 ggagtcgctt cgtagccaac gccgctccca ccatgcagcc tccggcggag tccagggtct 240
29 ccacggtagt cgtatgtcac ctgcgcgacc gtagctatcc gatctacatc ggcgcaggcc 300
30 tcctcgacga gccggacctg ctgcagaggg atgttcatgg taagagggtt ctggtggtga 360
31 ccaacaagac cgtgcgcgcg ctttacctgg acaagggtgac atgggcactc acccacaaca 420
32 acctgaatgt atcagtggaa agcgtgatcc tgcccagcgg tgaaaagtac aaaaatatgg 480
33 acacgctgat gaagggtgtt gacaaggcag tcgagtcccg ttttgaccgc cgggtgcacat 540
34 ttgtagcact ggggtggtgg gtcattgggg acatgtgtgg ttttgacgct gctgcattcc 600
35 tccggggcgt caatttcata cagataccaa ctactctgat ggcccagggt gattcatctg 660
36 ttggtgggaa aaccgggatt aaccacccac taggaaagaa cttgattggg gcattctacc 720
37 agccacaatg tgttttaatt gacacaaata cacttaacac attgcctgac agggagctag 780
38 cttcaggcat tgccgaggta gtaaagtatg ggctcataag ggatgcacca ttctttgagt 840
39 ggcaagagaa gaacatgccg aaattgttag caagagaacc aaatgctttg gcatatgcta 900
40 tcaagagatc atgtgaaaac aaagctgaag tgggtggcaca agatgagaag gaaagtggcc 960
41 ttcgagcaac actaaacctg ggtcacacat ttggccatgc tattgagact gggacaggat 1020
42 atggagcatg gctccatggg gaggtgtcgc cagctggaac agttatggca actgacatgt 1080
43 ctcaccgcct ggggtggata gatgactcca tcagaaaacg tgtggttgac atactaaagc 1140
44 aagccaaact tcccattgca cctcctgaga ccatgacogt agagaagttt aaaaacatca 1200
45 tggctgttga caagaagggt gctgatggtc tgttgagact catccttctg aaaggaccgc 1260
46 taggggtcgt tgtatttacg ggggattatg acgggaatgc actcgatgaa accctacatg 1320
47 cattctgcga caactgagac atcccagttt tggacatcac ttctgtatgc tagtactagg 1380
48 ctatgtatca atcagagaaa atattgtggg gtacattatc ttcactgcct cattggtata 1440
49 ttttttgtca gaaataatgg tatattgttt taagtgtctg tttgagtttg tagagaacca 1500
50 tggctaatac actgcggtca gttgttctct agtaataata gtaaagaagc ttcactacat 1560
51 cccttctgtc tgttactgtc aagtaacaat actcgtagag atttaaacca aaaaaaaaaa 1620
52 aaaaaaaaaa

```

54 <210> SEQ ID NO: 2

55 <211> LENGTH: 438

56 <212> TYPE: PRT

57 <213> ORGANISM: Zea mays

59 <400> SEQUENCE: 2

RAW SEQUENCE LISTING

DATE: 08/01/2001

PATENT APPLICATION: US/09/744,099

TIME: 18:19:59

Input Set : A:\BB1159 SEQ LST.txt

Output Set: N:\CRF3\08012001\I744099.raw

```

60 Met Ala Ala Ser Ala Ser Ser Leu Leu Ala Ala Pro Ala Ser Ser Ser
61   1           5           10           15
63 Cys Gly Ala Ile Ser Pro Gln Leu Pro Arg Gly Ala Pro Ala Ala Ala
64           20           25           30
66 Ser Val Ala Ser Pro Ser Arg His Ser Cys Tyr Leu Leu Arg Ala Ser
67           35           40           45
69 Pro Ser Arg Arg His Arg Ser Arg Phe Val Ala Asn Ala Ala Pro Thr
70           50           55           60
72 Met Gln Pro Pro Ala Glu Ser Arg Val Ser Thr Val Val Asp Val Asp
73 65           70           75           80
75 Leu Gly Asp Arg Ser Tyr Pro Ile Tyr Ile Gly Ala Gly Leu Leu Asp
76           85           90           95
78 Glu Pro Asp Leu Leu Gln Arg His Val His Gly Lys Arg Val Leu Val
79           100          105          110
81 Val Thr Asn Thr Thr Val Ala Pro Leu Tyr Leu Asp Lys Val Thr Trp
82           115          120          125
84 Ala Leu Thr His Asn Asn Leu Asn Val Ser Val Glu Ser Val Ile Leu
85           130          135          140
87 Pro Asp Gly Glu Lys Tyr Lys Asn Met Asp Thr Leu Met Lys Val Phe
88 145          150          155          160
90 Asp Lys Ala Val Glu Ser Arg Phe Asp Arg Arg Cys Thr Phe Val Ala
91           165          170          175
93 Leu Gly Gly Gly Val Ile Gly Asp Met Cys Gly Phe Ala Ala Ala Ala
94           180          185          190
96 Phe Leu Arg Gly Val Asn Phe Ile Gln Ile Pro Thr Thr Leu Met Ala
97           195          200          205
99 Gln Val Asp Ser Ser Val Gly Lys Thr Gly Ile Asn His Pro Leu
100          210          215          220
102 Gly Lys Asn Leu Ile Gly Ala Phe Tyr Gln Pro Gln Cys Val Leu Ile
103 225          230          235          240
105 Asp Thr Asn Thr Leu Asn Thr Leu Pro Asp Arg Glu Leu Ala Ser Gly
106           245          250          255
108 Ile Ala Glu Val Val Lys Tyr Gly Leu Ile Arg Asp Ala Pro Phe Phe
109           260          265          270
111 Glu Trp Gln Glu Lys Asn Met Pro Lys Leu Leu Ala Arg Glu Pro Asn
112           275          280          285
114 Ala Leu Ala Tyr Ala Ile Lys Arg Ser Cys Glu Asn Lys Ala Glu Val
115           290          295          300
117 Val Ala Gln Asp Glu Lys Glu Ser Gly Leu Arg Ala Thr Leu Asn Leu
118 305          310          315          320
120 Gly His Thr Phe Gly His Ala Ile Glu Thr Gly Thr Gly Tyr Gly Ala
121           325          330          335
123 Trp Leu His Gly Glu Ala Val Ala Ala Gly Thr Val Met Ala Thr Asp
124           340          345          350
126 Met Ser His Arg Leu Gly Trp Ile Asp Asp Ser Ile Arg Lys Arg Val
127           355          360          365
129 Val Asp Ile Leu Lys Gln Ala Lys Leu Pro Ile Ala Pro Pro Glu Thr
130           370          375          380
132 Met Thr Val Glu Lys Phe Lys Asn Ile Met Ala Val Asp Lys Lys Val

```

RAW SEQUENCE LISTING

DATE: 08/01/2001

PATENT APPLICATION: US/09/744,099

TIME: 18:19:59

Input Set : A:\BB1159 SEQ LST.txt

Output Set: N:\CRF3\08012001\I744099.raw

```

133 385          390          395          400
135 Ala Asp Gly Leu Leu Arg Leu Ile Leu Leu Lys Gly Pro Leu Gly Cys
136          405          410          415
138 Cys Val Phe Thr Gly Asp Tyr Asp Gly Asn Ala Leu Asp Glu Thr Leu
139          420          425          430
141 His Ala Phe Cys Asp Asn
142          435
144 <210> SEQ ID NO: 3
145 <211> LENGTH: 1358
146 <212> TYPE: DNA
147 <213> ORGANISM: Oryza sativa
149 <400> SEQUENCE: 3
150 gcacgagtac agcgtcgtct tcgtttcgca gtccaccac ctccatggcg gccgcgcct 60
151 cctcctctct gctcgccgcc gctcgtcgt cctcccgcgc ggcgccgtg tccgcccgcc 120
152 ggcgcgcag cgcgagtcgc gccgcggcgc cctccctccc ctgcgcgtct cgcgcctcct 180
153 ggcgtcctcc cctccgcgt tccgcagcga ggaccctccg cagccgtgtc gtcgcgagcg 240
154 cggccccgcg tatgcagccg ccgccggcgt cgagggtctc cagcgttgtc gacgtcgacc 300
155 tcggcgaccg gagctacccg atctacatcg gcgcaggcct tctcgacgag cctgacctgc 360
156 tgcagaggca tgttcatggg aagagggttt tgggtgtgac caacaccacc gtcgcgcgcg 420
157 tctacctgga gaaggtgacc tgggcactca cgcacaacaa cccgaatgtt tctgtggaga 480
158 gcgtgatcct gcccgacggc gagaagtaca aggacatggg cacctgatg aaggttttcg 540
159 acaaggcagt cgagtcgccg ctggaccggc ggtgcacgtt tgttgcggtg ggaggtggcg 600
160 ttattgggga catgtgcggt tttgcagctg ctgcgttctt gcgtggtgtc aatttcatac 660
161 agattcctac tactctgatg gcccggttg attcatctgt tggagggaag accggcatta 720
162 accatccatt ggggaagaac ttaattggg cgttctacca cccacagtgt gtactgatag 780
163 acaccgagac actgaataca ttgcctgaca gggaactggc ttcaggcata gctgaggtgg 840
164 tgaagtatgg tctcataaga gatgcaccgt tctttgaatg gcaagagaaa aacatgccag 900
165 cattattagc aagagaacca agtgctctgg cctatgctat taagagatcg tgtgaaaaca 960
166 aagctgaagt ggttgctcag gacgagaagg aaagtggctc ccgagcaaca cttaatctcg 1020
167 gccatacatt tggccatgct atagaaacag gaactggcta tggagcatgg ctccatgggg 1080
168 aggtctgttc agctggaaca gttatggcag ctgacatgtc tcaccgcctg ggttggatag 1140
169 acgagtcaat caagaaacgg gcaattgaca tactagagaa agcgaagctt ccaattacac 1200
170 ctccagaggg catgacagtg gagaagttca aaagtattat ggccgttgat aagaaggttg 1260
171 ctgatggatt gctgaggctc atcctcctga aaggacctct gggaagctgt gttttcactg 1320
172 gcgattactg ttcttcgagg tcgacgtgtc gataaatt 1358
174 <210> SEQ ID NO: 4
175 <211> LENGTH: 436
176 <212> TYPE: PRT
177 <213> ORGANISM: Oryza sativa
179 <400> SEQUENCE: 4
180 Met Ala Ala Ala Ala Ser Ser Ser Leu Leu Ala Ala Ala Ser Ser Ser
181 1 5 10 15
183 Ser Arg Ala Ala Val Ser Ala Arg Arg Ala Pro Ser Ala Ser Pro
184 20 25 30
186 Ala Ala Ala Ala Ser Leu Pro Ser Pro Ser Arg Ala Ser Cys Ala Pro
187 35 40 45
189 Pro Leu Arg Ala Ser Ala Ala Arg Thr Leu Arg Ser Arg Val Val Ala
190 50 55 60
192 Ser Ala Ala Pro Ala Met Gln Pro Pro Pro Ala Ser Arg Val Ser Thr

```

RAW SEQUENCE LISTING

DATE: 08/01/2001

PATENT APPLICATION: US/09/744,099

TIME: 18:19:59

Input Set : A:\BB1159 SEQ LST.txt

Output Set: N:\CRF3\08012001\I744099.raw

```

193 65          70          75          80
195 Val Val Asp Val Asp Leu Gly Asp Arg Ser Tyr Pro Ile Tyr Ile Gly
196          85          90          95
198 Ala Gly Leu Leu Asp Glu Pro Asp Leu Leu Gln Arg His Val His Gly
199          100          105          110
201 Lys Arg Val Leu Val Val Thr Asn Thr Thr Val Ala Pro Leu Tyr Leu
202          115          120          125
204 Glu Lys Val Thr Trp Ala Leu Thr His Asn Asn Pro Asn Val Ser Val
205          130          135          140
207 Glu Ser Val Ile Leu Pro Asp Gly Glu Lys Tyr Lys Asp Met Gly Thr
208 145          150          155          160
210 Leu Met Lys Val Phe Asp Lys Ala Val Glu Ser Arg Leu Asp Arg Arg
211          165          170          175
213 Cys Thr Phe Val Ala Leu Gly Gly Gly Val Ile Gly Asp Met Cys Gly
214          180          185          190
216 Phe Ala Ala Ala Ala Phe Leu Arg Gly Val Asn Phe Ile Gln Ile Pro
217          195          200          205
219 Thr Thr Leu Met Ala Gln Val Asp Ser Ser Val Gly Gly Lys Thr Gly
220          210          215          220
222 Ile Asn His Pro Leu Gly Lys Asn Leu Ile Gly Ala Phe Tyr His Pro
223 225          230          235          240
225 Gln Cys Val Leu Ile Asp Thr Glu Thr Leu Asn Thr Leu Pro Asp Arg
226          245          250          255
228 Glu Leu Ala Ser Gly Ile Ala Glu Val Val Lys Tyr Gly Leu Ile Arg
229          260          265          270
231 Asp Ala Pro Phe Phe Glu Trp Gln Glu Lys Asn Met Pro Ala Leu Leu
232          275          280          285
234 Ala Arg Glu Pro Ser Ala Leu Ala Tyr Ala Ile Lys Arg Ser Cys Glu
235          290          295          300
237 Asn Lys Ala Glu Val Val Ala Gln Asp Glu Lys Glu Ser Gly Leu Arg
238 305          310          315          320
240 Ala Thr Leu Asn Leu Gly His Thr Phe Gly His Ala Ile Glu Thr Gly
241          325          330          335
243 Thr Gly Tyr Gly Ala Trp Leu His Gly Glu Ala Val Ala Ala Gly Thr
244          340          345          350
246 Val Met Ala Ala Asp Met Ser His Arg Leu Gly Trp Ile Asp Glu Ser
247          355          360          365
249 Ile Lys Lys Arg Ala Ile Asp Ile Leu Glu Lys Ala Lys Leu Pro Ile
250          370          375          380
252 Thr Pro Pro Glu Ala Met Thr Val Glu Lys Phe Lys Ser Ile Met Ala
253 385          390          395          400
255 Val Asp Lys Lys Val Ala Asp Gly Leu Leu Arg Leu Ile Leu Leu Lys
256          405          410          415
258 Gly Pro Leu Gly Ser Cys Val Phe Thr Gly Asp Tyr Cys Ser Ser Arg
259          420          425          430
261 Ser Thr Cys Arg
262          435
264 <210> SEQ ID NO: 5
265 <211> LENGTH: 1643

```

RAW SEQUENCE LISTING

DATE: 08/01/2001

PATENT APPLICATION: US/09/744,099

TIME: 18:19:59

Input Set : A:\BB1159 SEQ LST.txt

Output Set: N:\CRF3\08012001\I744099.raw

266 <212> TYPE: DNA

267 <213> ORGANISM: Glycine max

269 <400> SEQUENCE: 5

```

270 gcacgagccc ggagtgggtt ttgtgttcaa ggaagacagg cgcaggcaca ggcacatcat    60
271 gttattgtta ttgttggctt cttcactcag caaccaccaa accacaccgt cgcacccatt    120
272 ccacgctga ttcccccaat ggcttccact gccaccaatt tctctcttcc tctctgcgcc    180
273 aaccaacaaa ctccaatccc caaaccctct ttcttctcca ataacaacca tttgcacttc    240
274 aactctaata ataattgggc ctgggcctct gtttccacct ctgcgaagtc aaggatatgc    300
275 gccacctcct ctcaagttat ggatccctcc gcagcaaaat ccgaaccgcg tcttcccacc    360
276 atcgtcgaag tcgatttggg tagccggagc tatcctattt acatcggatc cgggttacta    420
277 aaccaaccgc actatctcca gaggcattgt catggaaaga gagtctagt tgtaactaac    480
278 gaaaccgttg cgccacttta tctagacaag gttgttgatg ctttgacaag gggaaaccgc    540
279 aatgtttctg tggagagtgt aattttacct gatggtgagc agtacaagga catggatact    600
280 cttatgaaag tctttgacaa ggccatcgag tcgcggctcg accggcgggtg tacatttggt    660
281 gctcttggtg gtggtgtgat tggcgacatg tgtggctttg ctgcctctgc cttcctacgt    720
282 ggtgttaatt ttattcagat tcctacaact gtgatggcac aggtcgattc ttcagttggg    780
283 ggaaaaactg ggataaatca ccgccttggg aagaacatga tcggtacctt ttaccaacct    840
284 cagtgtgtgc ttatagacac agacacatta aatacgctac cggataggga actggcatca    900
285 gggctagcag aggttataaa gtatgggctc attagggatg cagagttttt tgagtggcaa    960
286 gagaaaaata tgcacttatt attggcaaga gatcctagt taatggcata tgctataaag   1020
287 cgatcttggtg aaaacaaggc tgaggttgtg tccttagatc agaaggaaaag tggactgagg   1080
288 gcaacattga acttgggtca tacatttggt catgcaatag aaactggggg aggtatggg   1140
289 cagtggcttc atggagaggc tgttgagcgt ggcacggtaa tggctgttga catgtcatat   1200
290 cgcctagggt ggattgatga ttctcttggt aaacgagttg gagacatttt aaaacaggct   1260
291 aagttaccca cagccctccc tgagaccgtg actgtggaca tgtttaaatc tgtcatggca   1320
292 gtggataaga aggtagcaga tgggttgcta aggttatcc ttctaaaggg tcctctaggc   1380
293 aattgtgttt tcacagggga ttatgacaga aaggctctag acaatacgtc tcgtgcattc   1440
294 tgtaaatoct gatctttgtg ctccatgttt ctattgcagc tttatccaat tccttttagt   1500
295 gagtttttgt atgttgttat ataacctcac ttgtattttt cttatgtact ttggatggca   1560
296 accatggttc ccattatata cgcaagagta tattatttga aattgaggca tgatcaataa   1620
297 ttattaaaaa aaaaaaaaaa aaa                                1643

```

299 <210> SEQ ID NO: 6

300 <211> LENGTH: 437

301 <212> TYPE: PRT

302 <213> ORGANISM: Glycine max

304 <400> SEQUENCE: 6

```

305 Met Ala Ser Thr Ala Thr Asn Phe Ser Leu Ser Leu Cys Ala Asn Gln
306   1             5             10             15
308 Gln Thr Pro Ile Pro Lys Pro Ser Phe Phe Ser Asn Asn Asn His Leu
309             20             25             30
311 His Phe Asn Ser Asn Asn Asn Trp Ala Trp Ala Ser Val Ser Thr Ser
312             35             40             45
314 Arg Lys Ser Arg Ile Cys Ala Thr Ser Ser Gln Val Met Asp Pro Ser
315             50             55             60
317 Ala Ala Lys Ser Glu Pro Ala Leu Pro Thr Ile Val Glu Val Asp Leu
318   65             70             75             80
320 Gly Ser Arg Ser Tyr Pro Ile Tyr Ile Gly Ser Gly Leu Leu Asn Gln
321             85             90             95
323 Pro Asp Tyr Leu Gln Arg His Val His Gly Lys Arg Val Leu Val Val

```


VERIFICATION SUMMARY

DATE: 08/01/2001

PATENT APPLICATION: US/09/744,099

TIME: 18:20:00

Input Set : A:\BB1159 SEQ LST.txt

Output Set: N:\CRF3\08012001\I744099.raw

L:9 M:270 C: Current Application Number differs, Replaced Application Number
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:13 M:256 W: Invalid Numeric Header Field, Wrong Prior FILING DATE:YYYY-MM-DD

POSTED: 08-01-2001

SEQUENCE LISTING

<110> E. I. du Pont de Nemours and Company

<120> Chorismate Biosynthesis Enzymes

<130> BB-1159

<140>

<141>

<150> 60/093,611

<151> July 21, 1998

<160> 9

<170> Microsoft Office 97

<210> 1

<211> 1628

<212> DNA

<213> Zea mays

<400> 1

```

ccgccgcttt cccgagctctc atggcggcgt ccgcgtcctc cctccttgcc gcccccgcat    60
cctcatcctg tggggccatc tccccccagc tccctcgtgg agctccggcc gctgcctccg   120
tcgcttcgcc gtcgcgtcac tccctgtacc tccctccggc tcccccttcg aggcgccatc   180
ggagtcgctt cgtagccaac gccgctccca ccatgcagcc tccggcggag tccagggtct   240
ccacggtagt cgatgtcgac ctccggcagc gtagctatcc gatctacatc ggcgcaggcc   300
tcctcgacga gccggacctg ctgcagagcc atgttcatgg taagagggtt ctggtggtga   360
ccaacacgac cgtcgcgcgc ctttacctgg acaagggtgac atgggcactc acccacaaca   420
acctgaatgt atcagtggaag agcgtgatcc tgcccgcagg tgaagagtag aaaaatatgg   480
acacgctgat gaagggtgtt gacaaggcag tcagagtcgg ttttgaccgc cgggtgcacat   540
ttgtagcact ggggtggtgg gtcatgtggg acatgtgtgg ttttgacagt gctgcattcc   600
tccggggcgt caatttcata cagataccaa ctactctgat ggcccagggt gattcatctg   660
ttggtgggaa aaccgggatt aaccacccac taggaaagaa cttgattggg gcattctacc   720
agccacaatg tgttctaatt gacacaaata cacttaacac attgcctgac agggagctag   780
cttcaggcat tgccgaggta gtaaaagtat ggctcataag ggatgcacca ttctttgagt   840
ggcaagagaa gaacatgccg aaattgttag caagagaacc aaatgctttg gcatatgcta   900
tcaagagatc atgtgaaaac aaagctgaag tgggtggcaca agatgagaag gaaagtggcc   960
ttcgagcaac actaaacctg ggtcacacat ttggccatgc tattgagact gggacaggat  1020
atggagcatg gctccatggg gaggtgtcgc cagctggaac agttatggca actgacatgt  1080
ctcaccgcct ggggtggata gatgactcca tcagaaaacg tgtgggttgac atactaaagc  1140
aagccaaact tcccattgca cctcctgaga ccatgaccgt agagaagttt aaaaacatca  1200
tggctgttga caagaagggt gctgatggtc tgttgagact catccttctg aaaggaccgc  1260
taggggtgctg tgtatttacg ggggattatg acgggaatgc actcgatgaa accctacatg  1320
cattctgcga caactgagac atcccagttt tggacatcac ttctgtatgc tagtactagg  1380
ctatgtatca atcagagaaa atatttgtgt gtacattatc ttcactgcct catttgtata  1440
ttttttgtca gaaataatgg tatattgttt taagtgtcgg tttgagtttg tagagaacca  1500
tggctaatac actgcggtca gttgttctct agtaataata gtaaagaagc ttcactacat  1560
cccttctgtc tgttactgtc aagtaacaat actcgtagag atttaaacca aaaaaaaaaa  1620
aaaaaaaaa                                     1628

```

<210> 2

<211> 438

<212> PRT

<213> Zea mays

0044099 01601

<400> 2

Met Ala Ala Ser Ala Ser Ser Leu Leu Ala Ala Pro Ala Ser Ser Ser
1 5 10 15
Cys Gly Ala Ile Ser Pro Gln Leu Pro Arg Gly Ala Pro Ala Ala Ala
20 25 30
Ser Val Ala Ser Pro Ser Arg His Ser Cys Tyr Leu Leu Arg Ala Ser
35 40 45
Pro Ser Arg Arg His Arg Ser Arg Phe Val Ala Asn Ala Ala Pro Thr
50 55 60
Met Gln Pro Pro Ala Glu Ser Arg Val Ser Thr Val Val Asp Val Asp
65 70 75 80
Leu Gly Asp Arg Ser Tyr Pro Ile Tyr Ile Gly Ala Gly Leu Leu Asp
85 90 95
Glu Pro Asp Leu Leu Gln Arg His Val His Gly Lys Arg Val Leu Val
100 105 110
Val Thr Asn Thr Thr Val Ala Pro Leu Tyr Leu Asp Lys Val Thr Trp
115 120 125
Ala Leu Thr His Asn Asn Leu Asn Val Ser Val Glu Ser Val Ile Leu
130 135 140
Pro Asp Gly Glu Lys Tyr Lys Asn Met Asp Thr Leu Met Lys Val Phe
145 150 155 160
Asp Lys Ala Val Glu Ser Arg Phe Asp Arg Arg Cys Thr Phe Val Ala
165 170 175
Leu Gly Gly Gly Val Ile Gly Asp Met Cys Gly Phe Ala Ala Ala Ala
180 185 190
Phe Leu Arg Gly Val Asn Phe Ile Gln Ile Pro Thr Thr Leu Met Ala
195 200 205
Gln Val Asp Ser Ser Val Gly Gly Lys Thr Gly Ile Asn His Pro Leu
210 215 220
Gly Lys Asn Leu Ile Gly Ala Phe Tyr Gln Pro Gln Cys Val Leu Ile
225 230 235 240
Asp Thr Asn Thr Leu Asn Thr Leu Pro Asp Arg Glu Leu Ala Ser Gly
245 250 255
Ile Ala Glu Val Val Lys Tyr Gly Leu Ile Arg Asp Ala Pro Phe Phe
260 265 270
Glu Trp Gln Glu Lys Asn Met Pro Lys Leu Leu Ala Arg Glu Pro Asn
275 280 285
Ala Leu Ala Tyr Ala Ile Lys Arg Ser Cys Glu Asn Lys Ala Glu Val
290 295 300
Val Ala Gln Asp Glu Lys Glu Ser Gly Leu Arg Ala Thr Leu Asn Leu
305 310 315 320

0944099.01501

Gly His Thr Phe Gly His Ala Ile Glu Thr Gly Thr Gly Tyr Gly Ala
325 330 335

Trp Leu His Gly Glu Ala Val Ala Ala Gly Thr Val Met Ala Thr Asp
340 345 350

Met Ser His Arg Leu Gly Trp Ile Asp Asp Ser Ile Arg Lys Arg Val
355 360 365

Val Asp Ile Leu Lys Gln Ala Lys Leu Pro Ile Ala Pro Pro Glu Thr
370 375 380

Met Thr Val Glu Lys Phe Lys Asn Ile Met Ala Val Asp Lys Lys Val
385 390 395 400

Ala Asp Gly Leu Leu Arg Leu Ile Leu Leu Lys Gly Pro Leu Gly Cys
405 410 415

Cys Val Phe Thr Gly Asp Tyr Asp Gly Asn Ala Leu Asp Glu Thr Leu
420 425 430

His Ala Phe Cys Asp Asn
435

<210> 3
<211> 1358
<212> DNA
<213> Oryza sativa

<400> 3

gcacgagtag	agcgtcgtct	tcgttttcgca	gtccaccac	ctccatggcg	gccgcccgcct	60
cctcctctct	gctcgccgcc	gcctcgtcgt	cctcccgcgc	ggcgcccgctg	tccgcccgcct	120
gcgccccag	cgcgagtcct	gccgcggcgg	cctcccctccc	ctcgccgtct	cgcgccctcct	180
gcgtcctctc	cctccgcgt	tccgcagcga	ggaccctccg	cagccgtgtc	gtcgcgagcg	240
cggccccgc	tatgcagccg	ccgcggcgt	cgaggggtctc	cacgggtgtc	gacgtcgacc	300
tcggcgaccg	gagctaccgc	atctacatcg	gcgcaggcct	tctcgacgag	cctgacctgc	360
tgcagaggca	tgttcattggg	aagaggggtt	tgggtgggtgac	caacaccacc	gtcgcgccgc	420
tctacctgga	gaaggtgacc	tgggcactca	cgcacaacaa	cccgaatgtt	tctgtggaga	480
gcgtgatcct	gcccgcagcg	gagaagtaca	aggacatggg	cacactgatg	aaggttttcg	540
acaaggcagt	cgagtcccgc	ctggaccggc	ggtgcacgtt	tgttgcggtg	ggaggtggcg	600
ttattgggga	catgtgcggt	tttgagctg	ctgcgttcct	gcgtgggtgtc	aatttcatac	660
agattcctac	tactctgatg	gcccaggtgg	attcatctgt	tggagggaag	accggcatta	720
accatccatt	ggggaagaac	ttaattgggg	cgtttctacca	cccacagtgt	gtactgatag	780
acaccgagac	actgaataca	ttgcctgaca	gggaactggc	ttcaggcata	gctgaggtgg	840
tgaagtattg	tctcataaga	gatgcaccgt	tctttgaatg	gcaagagaaa	aacatgccag	900
cattattagc	aagagaacca	agtgtctctg	cctatgctat	taagagatcg	tgtgaaaaca	960
aagctgaagt	ggttgctcag	gacgagaagg	aaagtggctc	ccgagcaaca	cttaattctcg	1020
gccatacatt	tggccatgct	atagaaacag	gaactggcta	tggagcatgg	ctccatgggg	1080
aggctgttgc	agctggaaca	gttatggcag	ctgacatgtc	tcaccgcctg	ggttggtatg	1140
acgagtcaat	caagaaacgg	gcaattgaca	tactagagaa	agcgaagctt	ccaattacac	1200
ctccagaggc	catgacagtg	gagaagttca	aaagtattat	ggccgttgat	aagaaggttg	1260
ctgatggatt	gctgaggctc	atcctcctga	aaggacctct	gggaagctgt	gttttctactg	1320
gcgattactg	ttcttcgagg	tcgacgtgtc	gataaatt			1358

<210> 4
<211> 436
<212> PRT
<213> Oryza sativa

<400> 4

Met Ala Ala Ala Ala Ser Ser Ser Leu Leu Ala Ala Ala Ser Ser Ser
1 5 10 15

Ser Arg Ala Ala Ala Val Ser Ala Arg Arg Ala Pro Ser Ala Ser Pro
20 25 30

Ala Ala Ala Ala Ser Leu Pro Ser Pro Ser Arg Ala Ser Cys Ala Pro
35 40 45

Pro Leu Arg Ala Ser Ala Ala Arg Thr Leu Arg Ser Arg Val Val Ala
50 55 60

Ser Ala Ala Pro Ala Met Gln Pro Pro Pro Ala Ser Arg Val Ser Thr
65 70 75 80

Val Val Asp Val Asp Leu Gly Asp Arg Ser Tyr Pro Ile Tyr Ile Gly
85 90 95

Ala Gly Leu Leu Asp Glu Pro Asp Leu Leu Gln Arg His Val His Gly
100 105 110

Lys Arg Val Leu Val Val Thr Asn Thr Thr Val Ala Pro Leu Tyr Leu
115 120 125

Glu Lys Val Thr Trp Ala Leu Thr His Asn Asn Pro Asn Val Ser Val
130 135 140

Glu Ser Val Ile Leu Pro Asp Gly Glu Lys Tyr Lys Asp Met Gly Thr
145 150 155 160

Leu Met Lys Val Phe Asp Lys Ala Val Glu Ser Arg Leu Asp Arg Arg
165 170 175

Cys Thr Phe Val Ala Leu Gly Gly Gly Val Ile Gly Asp Met Cys Gly
180 185 190

Phe Ala Ala Ala Ala Phe Leu Arg Gly Val Asn Phe Ile Gln Ile Pro
195 200 205

Thr Thr Leu Met Ala Gln Val Asp Ser Ser Val Gly Gly Lys Thr Gly
210 215 220

Ile Asn His Pro Leu Gly Lys Asn Leu Ile Gly Ala Phe Tyr His Pro
225 230 235 240

Gln Cys Val Leu Ile Asp Thr Glu Thr Leu Asn Thr Leu Pro Asp Arg
245 250 255

Glu Leu Ala Ser Gly Ile Ala Glu Val Val Lys Tyr Gly Leu Ile Arg
260 265 270

Asp Ala Pro Phe Phe Glu Trp Gln Glu Lys Asn Met Pro Ala Leu Leu
275 280 285

Ala Arg Glu Pro Ser Ala Leu Ala Tyr Ala Ile Lys Arg Ser Cys Glu
290 295 300

Asn Lys Ala Glu Val Val Ala Gln Asp Glu Lys Glu Ser Gly Leu Arg
305 310 315 320

0944099-01501

Ala Thr Leu Asn Leu Gly His Thr Phe Gly His Ala Ile Glu Thr Gly
325 330 335

Thr Gly Tyr Gly Ala Trp Leu His Gly Glu Ala Val Ala Ala Gly Thr
340 345 350

Val Met Ala Ala Asp Met Ser His Arg Leu Gly Trp Ile Asp Glu Ser
355 360 365

Ile Lys Lys Arg Ala Ile Asp Ile Leu Glu Lys Ala Lys Leu Pro Ile
370 375 380

Thr Pro Pro Glu Ala Met Thr Val Glu Lys Phe Lys Ser Ile Met Ala
385 390 395 400

Val Asp Lys Lys Val Ala Asp Gly Leu Leu Arg Leu Ile Leu Leu Lys
405 410 415

Gly Pro Leu Gly Ser Cys Val Phe Thr Gly Asp Tyr Cys Ser Ser Arg
420 425 430

Ser Thr Cys Arg
435

<210> 5
<211> 1643
<212> DNA
<213> Glycine max

<400> 5
gcacgagccc ggagtgggtgt ttgtgttcaa ggaagacagg cgcaggcaca ggcacatcat 60
gttattgtta ttgttggcct cttcactcag caaccaccaa accacaccgt cgcacccatt 120
ccatcgctga ttcccccaat ggcttccact gccaccaatt tctctcttct tctctgcgcc 180
aaccaacaaa ctccaatccc caaacccctct ttcttctcca ataacaacca tttgcacttc 240
aactctaata ataattgggc ctgggcctct gtttccacct ctgcaagtc aaggatatgc 300
gccacctcct ctcaagttat ggatccctcc gcagcaaaat ccgaaccgc tcttcccacc 360
atcgctcgaag tcgatttggg tagccggagc tatcctattt acatcggatc cgggttacta 420
aaccaacccg actatctcca gaggcattgt catggaaaga gagtcctagt tgtaactaac 480
gaaaccgttg cgccacttta tctagacaag gttgttgatg ctttgacaag gggaaaccg 540
aatgtttctg tggagagtgt aattttacct gatggtgagc agtacaagga catggatact 600
cttatgaaag tctttgacaa ggccatcgag tcgcggtctg accggcgggtg tacatttgtt 660
gctcttgggt gtggtgtgat tggcgacatg tgtggctttg ctgcctctgc cttcctacgt 720
ggtgttaatt ttattcagat toctacaact gtgatggcac aggtcgattc ttcagttggt 780
ggaaaaactg ggataaatca ccgccttggg aagaacatga tcggtacctt ttaccaacct 840
cagtgtgtgc ttatagacac agacacatta aatacgctac cggataggga actggcatca 900
gggctagcag aggttataaaa gtatgggctc attagggatg cagagttttt tgagtggcaa 960
gagaaaaata tgcaattatt attggcaaga gatcctagt taatggcata tgctataaag 1020
cgatcttgtg aaaacaaggc tgaggttgtg tccttagatc agaaggaaag tggactgagg 1080
gcaacattga acttgggtca tacatttgggt catgcaatag aaactggggt aggctatggg 1140
cagtggcttc atggagaggc tgttgacagt ggcacggtaa tggctgttga catgtcatat 1200
cgctaggtt ggattgatga ttctcttgtg aaacgagttg gagacatttt aaaacaggct 1260
aagttaccca cagcccctcc tgagaccgtg actgtggaca tgtttaaatc tgtcatggca 1320
gtggataaga aggtagcaga tgggttgcta aggccttatcc ttctaaaggg tcctctaggc 1380
aattgtgttt tcacagggga ttatgacaga aaggctctag acaatacgct tcgtgcattc 1440
tgtaaatcct gatctttgtg ctccatgttt ctattgcagc tttatccaat tccttttagt 1500
gagtttttgt atgttgttat ataacctcac ttgtattttt cttatgtact ttggatggca 1560
accatgggtc ccattatata cgcaagagta tattatttga aattgaggca tgatcaataa 1620
ttattaaaaa aaaaaaaaaa aaa 1643

<210> 6
 <211> 437
 <212> PRT
 <213> Glycine max

<400> 6

Met	Ala	Ser	Thr	Ala	Thr	Asn	Phe	Ser	Leu	Ser	Leu	Cys	Ala	Asn	Gln
1				5					10					15	
Gln	Thr	Pro	Ile	Pro	Lys	Pro	Ser	Phe	Phe	Ser	Asn	Asn	Asn	His	Leu
			20					25					30		
His	Phe	Asn	Ser	Asn	Asn	Asn	Trp	Ala	Trp	Ala	Ser	Val	Ser	Thr	Ser
		35					40					45			
Arg	Lys	Ser	Arg	Ile	Cys	Ala	Thr	Ser	Ser	Gln	Val	Met	Asp	Pro	Ser
	50					55					60				
Ala	Ala	Lys	Ser	Glu	Pro	Ala	Leu	Pro	Thr	Ile	Val	Glu	Val	Asp	Leu
	65				70					75					80
Gly	Ser	Arg	Ser	Tyr	Pro	Ile	Tyr	Ile	Gly	Ser	Gly	Leu	Leu	Asn	Gln
				85					90					95	
Pro	Asp	Tyr	Leu	Gln	Arg	His	Val	His	Gly	Lys	Arg	Val	Leu	Val	Val
			100					105					110		
Thr	Asn	Glu	Thr	Val	Ala	Pro	Leu	Tyr	Leu	Asp	Lys	Val	Val	Asp	Ala
		115					120					125			
Leu	Thr	Arg	Gly	Asn	Pro	Asn	Val	Ser	Val	Glu	Ser	Val	Ile	Leu	Pro
	130					135					140				
Asp	Gly	Glu	Gln	Tyr	Lys	Asp	Met	Asp	Thr	Leu	Met	Lys	Val	Phe	Asp
145					150					155					160
Lys	Ala	Ile	Glu	Ser	Arg	Leu	Asp	Arg	Arg	Cys	Thr	Phe	Val	Ala	Leu
				165					170					175	
Gly	Gly	Gly	Val	Ile	Gly	Asp	Met	Cys	Gly	Phe	Ala	Ala	Ser	Ala	Phe
			180					185					190		
Leu	Arg	Gly	Val	Asn	Phe	Ile	Gln	Ile	Pro	Thr	Thr	Val	Met	Ala	Gln
		195					200					205			
Val	Asp	Ser	Ser	Val	Gly	Gly	Lys	Thr	Gly	Ile	Asn	His	Arg	Leu	Gly
	210					215					220				
Lys	Asn	Met	Ile	Gly	Thr	Phe	Tyr	Gln	Pro	Gln	Cys	Val	Leu	Ile	Asp
225					230					235					240
Thr	Asp	Thr	Leu	Asn	Thr	Leu	Pro	Asp	Arg	Glu	Leu	Ala	Ser	Gly	Leu
				245					250					255	
Ala	Glu	Val	Ile	Lys	Tyr	Gly	Leu	Ile	Arg	Asp	Ala	Glu	Phe	Phe	Glu
			260					265					270		
Trp	Gln	Glu	Lys	Asn	Met	His	Leu	Leu	Leu	Ala	Arg	Asp	Pro	Ser	Val
		275					280					285			

09744099 "011601
 T09744099"

Met Ala Tyr Ala Ile Lys Arg Ser Cys Glu Asn Lys Ala Glu Val Val
 290 295 300

Ser Leu Asp Gln Lys Glu Ser Gly Leu Arg Ala Thr Leu Asn Leu Gly
 305 310 315 320

His Thr Phe Gly His Ala Ile Glu Thr Gly Val Gly Tyr Gly Gln Trp
 325 330 335

Leu His Gly Glu Ala Val Ala Ala Gly Thr Val Met Ala Val Asp Met
 340 345 350

Ser Tyr Arg Leu Gly Trp Ile Asp Asp Ser Leu Val Lys Arg Val Gly
 355 360 365

Asp Ile Leu Lys Gln Ala Lys Leu Pro Thr Ala Pro Pro Glu Thr Val
 370 375 380

Thr Val Asp Met Phe Lys Ser Val Met Ala Val Asp Lys Lys Val Ala
 385 390 395 400

Asp Gly Leu Leu Arg Leu Ile Leu Leu Lys Gly Pro Leu Gly Asn Cys
 405 410 415

Val Phe Thr Gly Asp Tyr Asp Arg Lys Ala Leu Asp Asn Thr Leu Arg
 420 425 430

Ala Phe Cys Lys Ser
 435

<210> 7
 <211> 1103
 <212> DNA
 <213> Triticum aestivum

<400> 7

gcacgagcat	tcggttgcgt	aggtggcgg	gtcattgggg	acatgtgtgg	gtttgcagct	60
gctgcattcc	tccgtgggtg	caatttcata	caaataccca	caactctgat	ggcccaggtg	120
gattcttctg	ttggagggaa	gactgggatt	aaccatccac	tgggaaagaa	cttaattggc	180
gcattctacc	agccgcagtg	tgtactcatt	gacacagaga	caactgaatac	attgcctgac	240
agagagctgg	cttcagggtg	cgctgaggtg	gtgaagtatg	gtctcataag	agatgcaccc	300
ttctttgagt	ggcaagaaaa	gaatatggca	gcaatcttag	cgagagaacc	aagtgccctg	360
acctacgcta	taaagagatc	atgtgaaaac	aaagctgaag	ttgttgctca	ggatgagaag	420
gaaagtggtc	ttcgagcaac	actaaacctg	ggtcatacat	ttgggtcatgc	tatagaaaca	480
gggcttggct	atggagaatg	gctccacggg	gaggctgttg	ctgctggaac	ggttatggca	540
gctgacatgt	cttaccgcct	gggctggata	gacgagtcca	tcaagaaacg	gacatttgac	600
atactagatc	aagccaagct	tcccgtgaca	tcaccaaagg	gcatgacggg	agagaagttc	660
agaaacatca	tggcgggtcg	caagaaggtc	gcggacggat	tggtgaggct	catccttctg	720
aaaggacctc	tgggaggctg	cgttttcacc	ggcgagtacg	acaggaaagc	cctggacgaa	780
accctccgog	ctttctgcga	caactgatcg	ttgtcaagct	ggagctccag	catttctgta	840
tactagtagg	ctgtaatatc	atatcagaga	gagagagaga	gagagtgcac	tgtgtgctac	900
ctttattttt	tcatttgcgt	ttggtgagaa	taatggaatg	cggttaagggg	ttggttccag	960
gtcgtagtac	taccaattca	ttgtagggca	tcattgcctca	gtaatcgtga	atcttgatca	1020
gtcattttga	gaaacacaat	tcggaacgag	atcctctgac	gtaggaagtt	ttaaaaaaag	1080
agaagaaaaa	aaaaaaaaaa	aaa				1103

<210> 8
 <211> 268
 <212> PRT
 <213> Triticum aestivum

<400> 8
 Ala Arg Ala Phe Val Ala Leu Gly Gly Gly Val Ile Gly Asp Met Cys
 1 5 10 15
 Gly Phe Ala Ala Ala Ala Phe Leu Arg Gly Val Asn Phe Ile Gln Ile
 20 25 30
 Pro Thr Thr Leu Met Ala Gln Val Asp Ser Ser Val Gly Gly Lys Thr
 35 40 45
 Gly Ile Asn His Pro Leu Gly Lys Asn Leu Ile Gly Ala Phe Tyr Gln
 50 55 60
 Pro Gln Cys Val Leu Ile Asp Thr Glu Thr Leu Asn Thr Leu Pro Asp
 65 70 75 80
 Arg Glu Leu Ala Ser Gly Val Ala Glu Val Val Lys Tyr Gly Leu Ile
 85 90 95
 Arg Asp Ala Pro Phe Phe Glu Trp Gln Glu Lys Asn Met Ala Ala Ile
 100 105 110
 Leu Ala Arg Glu Pro Ser Ala Leu Thr Tyr Ala Ile Lys Arg Ser Cys
 115 120 125
 Glu Asn Lys Ala Glu Val Val Ala Gln Asp Glu Lys Glu Ser Gly Leu
 130 135 140
 Arg Ala Thr Leu Asn Leu Gly His Thr Phe Gly His Ala Ile Glu Thr
 145 150 155 160
 Gly Leu Gly Tyr Gly Glu Trp Leu His Gly Glu Ala Val Ala Ala Gly
 165 170 175
 Thr Val Met Ala Ala Asp Met Ser Tyr Arg Leu Gly Trp Ile Asp Glu
 180 185 190
 Ser Ile Lys Lys Arg Thr Phe Asp Ile Leu Asp Gln Ala Lys Leu Pro
 195 200 205
 Val Thr Ser Pro Lys Gly Met Thr Val Glu Lys Phe Arg Asn Ile Met
 210 215 220
 Ala Val Asp Lys Lys Val Ala Asp Gly Leu Leu Arg Leu Ile Leu Leu
 225 230 235 240
 Lys Gly Pro Leu Gly Gly Cys Val Phe Thr Gly Glu Tyr Asp Arg Lys
 245 250 255
 Ala Leu Asp Glu Thr Leu Arg Ala Phe Cys Asp Asn
 260 265

<210> 9
 <211> 362
 <212> PRT
 <213> Escherichia coli

09744099-011501

<400> 9

Met Glu Arg Ile Val Val Thr Leu Gly Glu Arg Ser Tyr Pro Ile Thr
1 5 10 15
Ile Ala Ser Gly Leu Phe Asn Glu Pro Ala Ser Phe Leu Pro Leu Lys
20 25 30
Ser Gly Glu Gln Val Met Leu Val Thr Asn Glu Thr Leu Ala Pro Leu
35 40 45
Tyr Leu Asp Lys Val Arg Gly Val Leu Glu Gln Ala Gly Val Asn Val
50 55 60
Asp Ser Val Ile Leu Pro Asp Gly Glu Gln Tyr Lys Ser Leu Ala Val
65 70 75 80
Leu Asp Thr Val Phe Thr Ala Leu Leu Gln Lys Pro His Gly Arg Asp
85 90 95
Thr Thr Leu Val Ala Leu Gly Gly Gly Val Val Gly Asp Leu Thr Gly
100 105 110
Phe Ala Ala Ala Ser Tyr Gln Arg Gly Val Arg Phe Ile Gln Val Pro
115 120 125
Thr Thr Leu Leu Ser Gln Val Asp Ser Ser Val Gly Gly Lys Thr Ala
130 135 140
Val Asn His Pro Leu Gly Lys Asn Met Ile Gly Ala Phe Tyr Gln Pro
145 150 155 160
Ala Ser Val Val Val Asp Leu Asp Cys Leu Lys Thr Leu Pro Pro Arg
165 170 175
Glu Leu Ala Ser Gly Leu Ala Glu Val Ile Lys Tyr Gly Ile Ile Leu
180 185 190
Asp Gly Ala Phe Phe Asn Trp Leu Glu Glu Asn Leu Asp Ala Leu Leu
195 200 205
Arg Leu Asp Gly Pro Ala Met Ala Tyr Cys Ile Arg Arg Cys Cys Glu
210 215 220
Leu Lys Ala Glu Val Val Ala Ala Asp Glu Arg Glu Thr Gly Leu Arg
225 230 235 240
Ala Leu Leu Asn Leu Gly His Thr Phe Gly His Ala Ile Glu Ala Glu
245 250 255
Met Gly Tyr Gly Asn Trp Leu His Gly Glu Ala Val Ala Ala Gly Met
260 265 270
Val Met Ala Ala Arg Thr Ser Glu Arg Leu Gly Gln Phe Ser Ser Ala
275 280 285
Glu Thr Gln Arg Ile Ile Thr Leu Leu Lys Arg Ala Gly Leu Pro Val
290 295 300
Asn Gly Pro Arg Glu Met Ser Ala Gln Ala Tyr Leu Pro His Met Leu
305 310 315 320

09744099-011501

Arg	Asp	Lys	Lys	Val	Leu	Ala	Gly	Glu	Met	Arg	Leu	Ile	Leu	Pro	Leu
				325					330					335	
Ala	Ile	Gly	Lys	Ser	Glu	Val	Arg	Ser	Gly	Val	Ser	His	Glu	Leu	Val
			340					345					350		
Leu	Asn	Ala	Ile	Ala	Asp	Cys	Gln	Ser	Ala						
		355					360								

0944099.041604